Wolf Attack Probability: A New Security Measure in Biometric Authentication Systems

Masashi Une[1], Akira Otsuka[1], and Hideki Imai[1,2]

[1] Research Center for Information Security (RCIS),
National Institute of Advanced Industrial Science and Technology (AIST),
Akihabara-Daibiru Room 1102, 1-18-13, Sotokanda, Chiyoda, Tokyo 101-0021, Japan,
masashi-une@aist.go.jp,
WWW home page: http://www.rcis.aist.go.jp/
[2] Faculty of Science and Engineering, Chuo University,
1-13-27, Kasuga, Bunkyo, Tokyo 112-8551, Japan

Abstract. This paper will propose a wolf attack probability (WAP) as a new measure for evaluating security of biometric authentication systems. The wolf attack is an attempt to impersonate a victim by feeding "wolves" into the system to be attacked. The "wolf" means an input value which can be falsely accepted as a match with multiple templates. WAP is defined as a maximum success probability of the wolf attack with one wolf sample. In this paper, we give a rigorous definition of the new security measure which gives strengh estimation of an individual biometric authentication system against impersonation attacks. We show that if one reestimates using our WAP measure, a typical fingerprint algorithm is turned out to be much weaker than theoretically estimated by Ratha et al. Moreover, we apply the wolf attack to a finger-vein-pattern matching algorithm. Surprisingly, we show that there exists an extremely strong wolf which falsely matches all templates for any threshold values.

1 Introduction

A biometric authentication system automatically authenticates an individual by using physiological and/or behavioral characteristics. Recently, the use of biometric authentication systems has spread in various services such as the immigration control at an airport, financial transactions at an ATM (automated tellers machine) terminal, the access control for a mobile phone, and so on. This trend has made it more important to exactly evaluate the security of biometric authetication systems.

In order to conduct the security evaluation, it is necessary to identify threats and vulnerabilities regarding the biometrics. General threats and vulnerabilities common to biometric authentication systems have already been clarified in many literatures. The committee draft of ISO/IEC 19792 describes three threats and eleven vulnerabilities [1]. With regard to the threats, the draft includes "intentional impersonation," "unexpected high FAR (false acceptance rate)," and "creating backdoor." Focusing on the intentional impersonation, the following

three attacks have been widely discussed so far: a brute-force attack, a zero-effort attack, and an artifact attack. The artifact attack is that an attacker presents the victim's biometric characteristic by using some artifacts[2]. Although a security measure against the zero-effort attack has been established as FAR, no measures are commonly accepted by the industry against the brute-force attack and the artifact attack.

With regard to the brute-force attack, Ratha et al.[3] estimates its success probability in a fingerprint-minutiae matching algorithm. We will call the success probability a "minutiae collision probability (MCP)." However, MCP of Ratha et al. is computed under the following condition: the attacker presents an input value that consists of the same number of minutiae as that of all templates to be compared with. Therefore, MCP of Ratha et al. does not give the exact success probability of the brute-force attack. In order to do so, we have to compute MCP in such a way to take all possible input values into account.

The brute-force attack is supposed to be carried out under the situation that an attacker blindly selects an input value to be presented to a biometric authentication system. However, if we assume that the attacker has some information on the internal algorithms employed in the system, we have to pay attention to an attack with a smarter choice of a sample, "wolf." The draft of ISO/IEC 19792 defines the wolf as a biometric sample that shows high similarity to most of the templates[1]. If the attacker successfully found the wolf, he could impersonate the victim with a higher probability than MCP by presenting the wolf.

In this paper, we call such an attack a "wolf attack," and propose a "wolf attack probability (WAP)" as a maximum success probability of the wolf attack with one wolf sample. WAP is considered to be the upper bound of the success probability of attacks that are carried out without knowledge of a victim's biometric sample. Therefore, WAP can be used as a security measure to evaluate the lower bound of a security level in an individual biometric authentication system.

We show that WAP is extremely larger than the theoretical estimation of MCP by Ratha et al. in the fingerprint-minutiae matching algorithm. Ratha et al. computed MCP under the condition that the number of minutiae in the input value N_q is identical to that of minutiae in the template N_r. Especially, they discussed MCP for $N_q = N_r = 40$. On the other hand, WAP is given as the maximum of MCP where $N_q = 400$ and $N_r = 40$. For example, while Ratha et al. obtained $MCP = 2^{-80}$ for threshold $m = 25$, we show that $WAP = 2^{-20}$. In this case, we can understand that the wolf attack gains the attack complexity of about 2^{60}.

Moreover, we will apply the wolf attack to a finger-vein-pattern matching alorithm proposed by [4]. Surprisingly, we show that there exists the wolf which falsely matches any templates for any threshold values. Especially, we call such a wolf a "universal wolf." This result implies that it is necessary to evaluate an impact of the wolf attack on a matching algorithm by applying the wolf attack and obtaining WAP.

This paper continues as follows. Section 2 will obtain exact MCP on the basis of [3]. By using MCP, we will search for an input value that maximizes the probability of a false match with a given template, and show that the maximized probability is extremely larger than MCP of Ratha et al. Section 3 will define the wolf attack and WAP. Section 4 will describe FAR and discuss its limitation as a security measure for the wolf attack. Section 5 will show a result of applying the wolf attack to the finger-vein-pattern matching alorithm proposed by [4]. Section 6 will summarize our results and show future research topics.

2 Brute-Force Attack in a Fingerprint-Minutiae Matching Algorithm

2.1 Minutiae collision probability by Ratha et al.

Ratha et al.[3] discusses a typical fingerprint-minutiae matching algorithm in which the number of matched minutiae between an input value and a template reflects the degree of the match. The feature of a minutia consists of its location (x, y) and the ridge direction d. If the number of paired minutiae whose locations and ridge directions are equal to or more than a threshold value m, the input value is accepted as a match with the corresponding template.

In such a matching algorithm, Ratha et al. discussed the security level against the brute-force attack. Assuming that the attacker presents an input value consisting of forged minutiae whose locations and ridge directions are randomly selected, Ratha et al. attempted to compute the probability that the input value falsely matches a given template in both a location and a ridge direction. We call the probability a "minutiae collision probability (MCP)." In general, the definition of MCP is given as follows.

Definition 1. *Let S_{N_q} and T_{N_r} be a set of input values consisting of N_q minutiae and a set of templates consisting of N_r minutiae, respectively. Let match be a function that has two inputs $s(\in S_{N_q})$ and $t(\in T_{N_r})$ and an output of "accept" or "reject" as the result of the match. For given N_q and N_r,*

$$MCP \stackrel{\triangle}{=} \underset{s \in S_{N_q}}{Ave} \underset{t \in T_{N_r}}{Ave} Pr[match(s,t) = accept] \tag{1}$$

where Pr[X] and Ave Y denote a probability of the occurrence of phenomenon X and a mean of Y, respectively.

In [3], MCP is discussed under the condition of $N_q = N_r$. This condition means that the number of minutiae in the input value is identical to that of the template. As a result, MCP_R, which denotes conditional MCP computed by Ratha et al., is defined as follows.

Definition 2. *For a given $N_p(= N_q = N_r)$,*

$$MCP_R \stackrel{\triangle}{=} \underset{s \in S_{N_p}}{Ave} \underset{t \in T_{N_p}}{Ave} Pr[match(s,t) = accept]. \tag{2}$$

The difference between MCP and MCP_R is whether the condition of $N_q = N_r$ is applied or not.

On the basis of the definition of MCP_R, Ratha et al. employs the following p_{hi} as a probability that a minutia selected randomly is included in the template:

$$p_{hi} = \frac{N_p}{(K - N_p + 1)d} \tag{3}$$

where K denotes the number of possible minutiae locations. This value p_{hi} indicates a probability that after $N_p - 1$ minutiae fail to match, the N_pth minutia matches. Therefore, p_{hi} is the conservative approximation of the probability.

As a result, Ratha et al. obtained MCP_R as follows:

$$MCP_R = \sum_{t=m}^{N_p} \binom{N_p}{t} (p_{hi})^t (1 - p_{hi})^{N_p - t}. \tag{4}$$

2.2 Computing MCP

Let us consider the condition of $N_q = N_r$ in MCP_R. With regard to this condition, Ratha et al. describes as follows: "Note that brute force attacks with N_q excessively large (close to the value K) would be easy to detect and reject out of hand." This claim may be correct when considering not only the security level of the matching algorithm itself but also some additional countermeasures that reduce the strength of the brute-force attack. However, in order to focus on the security level of the algorithm itself as an objective to be evaluated, we should at first discuss MCP without putting any conditions on N_q and N_r.

To handle with the various values of N_q and N_r, we will give the exact probability of MCP instead of relying on the approximation used by Ratha et al.

Let us compute MCP instead of MCP_R. At first, we obtain probability P_N that $N(\geq m)$ of N_q minutiae in an input value match with regard to their locations. Then, we obtain probability P'_N that m of the N minutiae match with regard to their ridge directions. As a result, MCP is expressed as follows:

$$MCP = \sum_{N=m}^{N_r} (P_N \times P'_N). \tag{5}$$

P_N and P'_N are expressed as follows.

$$P_N = \frac{\binom{N_q}{N}\binom{K - N_q}{N_r - N}}{\binom{K}{N_r}}. \tag{6}$$

$$P'_N = \sum_{t=m}^{N} \binom{N}{t} \left(\frac{1}{d}\right)^t \left(1 - \frac{1}{d}\right)^{N-t}. \tag{7}$$

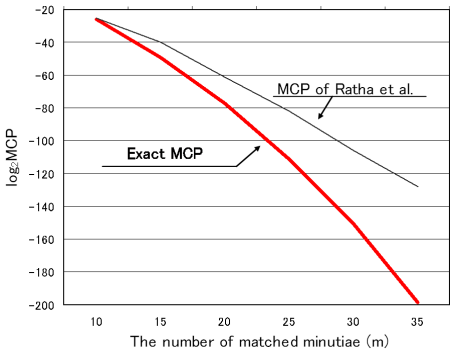

Fig. 1. Comparison between MCP and MCP_R for $N_q = N_r = 40$

2.3 Comparison between MCP and MCP_R

Let us calculate concrete values of MCP and MCP_R by using equations (4)-(7). Basically, we employ the following parameters used by Ratha et al.

- The size of an input value $S = 300 \times 300$ pixels.
- A ridge plus valley spread $T = 15$ pixels.
- The total number of possible minutiae sites $(K = S/(T^2)) = 20 \times 20 = 400$.
- The number of orientations allowed for the ridge angle at a minutia point $d = 4$.
- The minimum number of corresponding minutiae in an input value and template, i.e., a threshold value $m = 10, 15, 20, 25, 30, 35$.

At first, let us compare exact MCP with MCP_R for $N_q = N_r = 40$ (see fig.1). Figure 1 indicates that MCP_R is larger than MCP. For example, while MCP_R is about 2^{-80} for $m = 25$, MCP is about 2^{-111}. As mentioned above, Ratha et al. approximated the success probability of the brute-force attack in the conservative manner, thus they slightly overestimated the strength of the brute-force attack.

Next, let us calculate MCP for any N_q and m. N_r is fixed as 40 as previous. The result is shown in fig. 2. As expected by Ratha et al., figure 2 indicates that the input values of $N_q = 400$ make MCP maximized for any m. In case of $m = 25$, MCP is maximized as about 2^{-20} for $N_q = 400$ and as about 2^{-111} for $N_q = 40$, respectively. It turns out that the attacks using the fingerprints with 400 minutiae extremely gains the attack complexity of about 2^{90} in comparison with the attacks using the fingerprints with only 40 minutiae.

3 Wolf Attack and Wolf Attack Probability

A phenomenon that a special input value causes an extremely high success probability of the false match has been recognized mainly in the field of the speaker

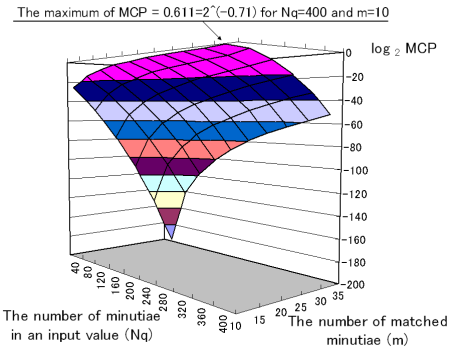

Fig. 2. MCP for $N_r = 40$

recognition[5]. Such an input value is called a *wolf*. As shown in [1], it has been common that the wolf means a biometric sample, not a synthesised one.

However, considering the case of input values with 400 minutiae described at the previous section, we should assume that an attacker may successfully find the special input value not only from a group of biometric samples but also from that of non-biometric samples. Moreover, we also have to pay attention to the fact that some biometric authentication systems falsely accept non-biometric samples presented by artifacts [2]. When one wants to evaluate the security of a certain biometric authentication system, we should take it into account that the attacker successfully finds some special input value, "wolf," and may present it by using some artifacts.

We define the wolf as follows.

Definition 3. *Let S_A be a group of all possible input values including ones taken from artifacts. Let T_h be a group of templates taken from all human samples. A wolf is defined as an input value $s_w \in S_A$ such that $match(s_w, t) = accept$ for multiple templates $t \in T_h$.*

We call an input value, s_w, is "p-wolf," if the matching probability of the input value s_w is given by $p = Ave_{t \in T_h} Pr[match(s_w, t) = accept]$. Especially, we call *1*-wolf a "universal wolf," which falsely matches any templates with probability $p = 1$.

We also define the wolf attack and the wolf attack probability as follows.

Definition 4. *Assume that the following two conditions are satisfied. The one is that the attacker has no information of a biometric sample of a victim to be impersonated. The other is that the attacker has complete information of a matching algorithm in the biometric authentication system to be attacked.*

The wolf attack is defined as an attempt to impersonate the victim in such a way to present p-wolves with large p's to minimize the complexity of the impersonation attack.

Definition 5. *We define the wolf attack probability (WAP) as follows:*

$$WAP \stackrel{\triangle}{=} \max_{s_w \in S_A} \underset{t \in T_h}{Ave}\, Pr[match(s_w, t) = accept] \tag{8}$$

where $\max Z$ *denotes a maximum of* Z*.*

In the fingerprint-minutiae matching algorithm discussed at the previous section, the attack of presenting the forged fingerprint with 400 minutiae corresponds to the wolf attack. WAP is given as the maximum of MCP where $N_q = 400$ and $N_r = 40$.

Next, we will discuss the lower bound on the number of wolves to be presented in order to make any templates falsely matched under a given WAP.

Theorem 1. *Given a biometric authentication system with wolf attack probability* WAP*, suppose there exists an attacker who has a set of wolves* $\{s_{w_1}, s_{w_2}, \ldots, s_{w_q}\}$ *which covers the whole group of templates* T_h*. Then, the following equation holds:*

$$q \geq \frac{1}{WAP}. \tag{9}$$

Proof. Let $T_h^{w_i}$ be a group of templates falsely matched by s_{w_i} for $i = 1, 2, \ldots, q$. Since both $T_h = (T_h^{w_1} \cup T_h^{w_2} \cup \cdots \cup T_h^{w_q})$ and $|T_h^{w_i}|/|T_h| \leq WAP$ hold,

$$|T_h| = |T_h^{w_1} \cup T_h^{w_2} \cup \cdots \cup T_h^{w_q}| \leq \sum_{i=1}^{q} |T_h^{w_i}| \leq q \times |T_h| \times WAP. \tag{10}$$

Therefore, $q \geq 1/WAP$. □

The equality, $q = 1/WAP$, holds, when the following conditions are satisfied. The first is $T_h^{w_i} \cap T_h^{w_j} = \emptyset$ for all $i \neq j$. The second is $|T_h^{w_i}|/|T_h| = WAP$ for all i.

Thus, even though the attacker has a set of the wolves that covers all of the templates, the number of attempts required for falsely matches with any templates are lower bounded by $1/WAP$. Thus, WAP gives a good security measure for evaluating an individual biometric authentication system against general impersonation attacks considering the existence of wolves.

4 Is *FAR* Suitable for the Security Measure regarding the Wolf Attack?

In the previous section, we defined the wolf attack and WAP, and explained why we had to pay attention to them. Then, let us discuss whether or not FAR is suitable for the evaluation of a security level against the wolf attack, instead of WAP.

By using the terms of the definition of the wolf attack, we can define FAR as follows.

Definition 6. *Let $S_h(\subset S_A)$ be a group of input values taken from all human samples. Let T_h be a group of templates taken from all human samples. FAR is defined as the following probability:*

$$FAR \stackrel{\triangle}{=} \underset{s \in S_h}{Ave} \underset{t \in T_h}{Ave} Pr[match(s,t) = accept]. \tag{11}$$

Note that FAR is defined by using both biometric samples and templates from S_h and T_h, respectively. If both $S_h = S_{N_q}$ and $T_h = T_{N_r}$ hold in the fingerprint-minutiae matching algorithm, FAR is equal to MCP.

With regard to a relationship between FAR and WAP, we can obtain the following lemma.

Lemma 1. $FAR \leq WAP$.

Proof. Trivial. □

As discussed later, some biometric authentication system may have strong wolves, and in the extreme case the system may contain a universal wolf, hence $WAP = 1 \gg FAR$. If only FAR is given as the impersonation measure for a biometric authentication system, such a strong wolf is not explicitly indicated in the specification of the system.

5 Applying the Wolf Attack to a Finger-Vein-Pattern Matching Algorithm

We will demonstrate how to search for the wolf and obtain WAP by applying to a finger-vein-pattern matching algorithm proposed by [4].

5.1 Overview of the algorithm to be analyzed

Let us briefly introduce an overview of the algorithm proposed by [4]. A finger vein pattern, which is a binarized image of 240×180 pixels, is extracted from an infrared image of the finger. In generating an input value to the matching algorithm from the binarized image, spatial reduction and relabeling of pixels are performed.

In the spatial reduction, the binarized image is reduced to one third of its original size in x and y dimensions. In this process, the binarized image is divided into $4,800$ windows of 3×3 pixels, and a mean of the grayscale of each window is calculated. The grayscale of each pixel in the reduced image is assigned with the mean.

In the relabeling of the reduced image, which is used as the input value, each pixel in the reduced image is classified into the following three: a vein region, a background region and an ambiguous region. Pixels whose grayscales are between 171 and 255 are labeled as the vein region. Pixels whose grayscales are between 0 and 84 are labeled as the background region. The other pixels are labeled as

the umbiguous region. Then, grayscales of pixels in background, ambiguous and vein regions are reassigned with 0, 128 and 255, respectively.

The input value is matched with the corresponding template. The differentiation between the input value and the template is represented by a "mismatch ratio (R_m)." R_m is defined as follows.

Definition 7. *Let an input value and a template be*

$$I = \{x_{i,j} | x_{i,j} \in \{0, 128, 255\}, i = 1, 2, \ldots, 80, j = 1, 2, \ldots, 60\}, \textit{and} \quad (12)$$

$$T = \{y_{i,j} | y_{i,j} \in \{0, 128, 255\}, i = 1, 2, \ldots, 80, j = 1, 2, \ldots, 60\}, \quad (13)$$

respectively. $x_{i,j}$ *and* $y_{i,j}$ *denote grayscales of pixels at* (i, j) *in the input value and the template, respectively.* $x_{i,j} = 0, 128$ *and* 255 *indicate that a pixel at* (i, j) *in* I *belongs to background, ambiguous and vein regions, respectively.* $y_{i,j} = 0, 128$ *and* 255 *indicate that a pixel at* (i, j) *in* T *belongs to background, ambiguous and vein regions, respectively.*

R_m *is defined as follows:*

$$R_m = \frac{|\{x_{i,j} | |x_{i,j} - y_{i,j}| = 255\}|}{|\{x_{i,j} | x_{i,j} = 255\}| + |\{y_{i,j} | y_{i,j} = 255\}|}. \quad (14)$$

If R_m is equal to or less than the predetermined threshold, the input value is decided to match the template[1].

5.2 Searching a universal wolf

We show that there exists a universal wolf in the finger-vein-pattern matching algorithm as follows.

Lemma 2. *There exists a universal wolf* s_{uw} *in the finger-vein-pattern matching algorithm[4].*

Proof. In order to prove this lemma, it is sufficient to show one example of s_{uw} that causes $R_m = 0$ for any templates.

From the definition of R_m, it is clear that if all of pixels in the input value belong to the ambiguous region, $R_m = 0$ holds. Namely, if $x_{i,j} = 128$ for all (i, j) in s_{uw}, $|\{x_{i,j} | |x_{i,j} - y_{i,j}| = 255\}| = 0$ holds because there exists no $y_{i,j}$ such that $|128 - y_{i,j}| = 255$. Then, the following equation holds for any templates:

$$R_m = \frac{0}{0 + |\{y_{i,j} | y_{i,j} = 255\}|} = 0. \quad (15)$$

Thus, an input value in which all pixels belong to the ambiguous region is one example of s_{uw}. □

[1] [4] does not describe how to compute R_m if both a denominator and a numerator of R_m become zero. In this paper, we regard such R_m as zero.

We put emphasis on that the universal wolf exists in their "matching algorithm" only. Moreover, we add that we have not found a universal wolf for the whole biometric authentication system described in [4] including their probabilistic feature extraction process. In the real evaluation of WAP on an individual biometric authentication system, we need to analyze the feature extraction algorithm altogether.

6 Concluding Remarks

In this paper, we proposed the wolf attack and the wolf attack probability (WAP). We showed that WAP is extremely larger than the success probability which Ratha et al. estimated with regard to the brute-force attack in the fingerprint-minutiae matching algorithm. We also found the universal wolf in the finger-vein-pattern matching algorithm [4]. The universal wolf is an input value that falsely matches any templates for any threshold values.

We have proposed to evaluate the security level against the wolf attack by computing WAP. Assuming that the attacker attempts to impersonate a victim without the knowledge of the victim's biometric sample, WAP gives the lower bound of the security level satisfied by a biometric authentication system.

One of future research topics is to apply the wolf attack to the other matching algorithms. Such research results are useful in comparing the algorithms from the viewpoint of the security against the impersonation attack. Furthermore, one may think of "provable security" in the sense of non-existence of strong p-wolf for any $p > k_0$, with some security parameter k_0. It is an open problem to show a construction of provably-secure biometric authentication systems in the above sense.

References

1. International Organization for Standardization (ISO), International Electrotechnical Commission (IEC): ISO/IEC CD 19792: Information technology – Security techniques – Security evaluation of biometrics. (2006)
2. Matsumoto, T., Matsumoto, H., Yamada, K., Hoshino, S.,: Impact of Artificial 'Gummy' Fingers on Fingerprint Systems. Optical Security and Counterfeit Deterrence Techniques IV, Proceedings of SPIE **4677** (2002) 275-289
3. Ratha, N. K., Connell, J. H., Bolle R. M.,: Enhancing security and privacy in biometrics-based authentication systems. IBM Syst. J. **40** (2001) 614–634
4. Miura, N., Nagasaka, A., Miyatake, M.,: Feature Extraction of Finger Vein Patterns Based on Iterative Line Tracking and its Application to Personal Identification. IEICE Trans. Inf. & Syst. (Japanese Edition) **J-86-D-II** (2003) 678–687
5. Doddington, G., Liggett, W., Martin, A., Przybocki, M., Reynolds, D.,: SHEEP, GOATS, LAMBS and WOLVES: A Statistical Analysis of Speaker Performance in the NIST 1998 Speaker Recognition Evaluation. Proc. ICSLP 98 (1998) 1351–1354