BEAT

Biometrics Evaluation and Testing

http://www.beat-eu.org/

Funded under the 7th FP (Seventh Framework Programme)

Theme SEC-2011.5.1-1

[Evaluation of identification technologies, including Biometrics]

D4.5: Description of metrics for the evaluation of vulnerabilities to direct attacks

Due date: 30/04/2014
Project start date: 01/03/2012
WP Manager: Nils Tekampe, TUViT

Submission date: 17/04/2014
Duration: 48 months
Revision: 2

Author(s): Nils Tekampe (TUVIT), Alain Merle (CEA), Ivana Chingovska (Idiap), André Anjos (Idiap), Sébastien Marcel (Idiap)

Project funded by the European Commission in the 7th Framework Programme (2008-2010)		
Dissemination Level		
PU	Public	Yes
RE	Restricted to a group specified by the consortium (includes Commission Services)	No
CO	Confidential, only for members of the consortium (includes Commission Services)	No

D4.5: Description of metrics for the evaluation of vulnerabilities to direct attacks

Abstract:

The objective of D4.5 is to describe metrics for the evaluation of vulnerabilities to direct attacks. This document shares two different views for which ongoing harmonization efforts are still ongoing both at the European standardization level (CEN TC224 WG18) and at the International standardization level (ISO SC37 WD30107). On one hand the document proposes a metric to evaluate jointly the performance and the vulnerability to spoofing of biometric sub-system components (algorithmic level). On the other hand the document proposes as well a metric for the rating of direct (spoofing) attacks against biometric systems. Authors consider that both views are complementary and are currently proposing it at the ISO level for inclusion in the ISO Working Draft WD30107 on spoofing detection.

Contents

1 Introduction

Metrics for direct attacks against biometric systems are two-fold. On the one hand the performance of recognizing a spoofing attack as an attack can be described using a purely descriptive approach. Such an approach would be an equivalent to the error rates of a biometric system regarding its recognition performance. On the other hand, spoofing attacks are attacks against biometric systems aiming to break the core biometric functionality of the system. As such those aspects will need a metric to describe, how likely it is that a biometric system can be broken by an attacker. Both aspects are legitimate and cannot be seen in isolation. In order to decide about whether a biometric system is able to withstand a spoofing attack it is essential to know about the amount of fakes that is correctly recognized by the system. On the other hand the information that a biometric system will recognize 99 percent of all fakes does not allow the conclusion that this system can withstand an attack as there may be the one "'golden fake"' that reproducibly breaks the system.

In order to reflect both perspectives this document splits the proposed metric into two parts: Chapter 2 will introduce a metric to describe the performance of biometric systems in recognizing spoofing attacks. Chapter 3 will introduce a metric focusing on vulnerability assessment in context of Common Criteria.

Various documents have been considered for this metric. Namely

- Common Criteria for Information Security Evaluation ([3]),
- Characterizing Attacks to Fingerprint verification mechanisms ([9]),
- BEAT D4.6 - Description of Metrics for the evaluation of Vulnerabilities to indirect Attacks

The results documented in the present document will be considered for deliverable 6.5 forming a comprehensive evaluation methodology for biometric systems based on the Common Criteria. In D6.5 the various aspects of the metric described in this document will be linked to numeric values allowing to calculate a single numeric value expressing the strength of an attack. D6.5 will also propose a way to link this numeric value to the various levels of resistance defined in [3].

2 Performance Metric

A biometric recognition system under spoofing attacks can receive three types of inputs: genuine users, zero-effort impostors and spoofing attacks. Still, the system can keep its binary nature, as there is only one positive class of samples that need to be accepted (genuine users), while the samples that need to be rejected (zero-effort impostors and spoofing attacks) can be considered as one negative super-class.

Evaluating a biometric recognition system under spoofing attacks poses three important requirements:

- Defining a metric for assessing vulnerability to spoofing;
- Defining a single metric for assessing the overall performance of the system (both recognition performance and vulnerability to spoofing);
- Determining decision threshold for the system.

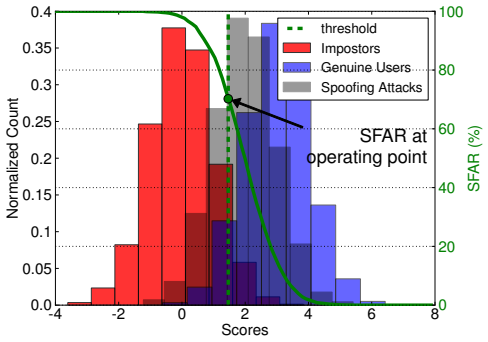
Subsections 2.1, 2.2 and 2.3 concern with each of these requirements, respectively.

2.1 Assessing the vulnerability to spoofing

The metric for assessing vulnerability to spoofing needs to be analogous to the well established metrics for assessing the recognition performance of the system: False Acceptance Rate (FAR), and False Rejection Rate (FRR). If FAR corresponds to the ratio of incorrectly accepted zero-effort impostors and FRR corresponds to the ratio of incorrectly rejected genuine users, then Spoof False Acceptance Rate (SFAR) is defined as the ratio of incorrectly accepted spoofing attacks [7].

An intuition of the vulnerability to spoofing of a biometric system can be obtained using a score distribution histogram. An example for a hypothetical system is given in Figure 1. The histogram shows the distribution of the output scores that the system gives for the three input classes: real accesses, zero-effort impostors and spoofing attacks. The green dotted line shows the threshold of the system. All the samples which are on the right of this line are accepted by the system. If they belong to the class of zero-effort impostors or spoofing attacks, then they contribute to FAR or SFAR, respectively. All the samples on the left of this line are rejected by the system. If they belong to the class of real samples, then they contribute to FRR. By moving the threshold to the right, FAR and SFAR will decrease, while FRR will increase. The full green curve shows how the spoofing vulnerability of the system decreases as the threshold is moved to the right.

Figure 1: Score distribution plot for biometric recognition system (hypothetical case)



2.2 Assessing the overall performance

To establish a metric that will jointly report on the recognition performance and vulnerability to spoofing of the system, a way to group the zero-effort impostors and spoofing attacks into the negative super-class is needed. The most straight-forward way to do it is simply to merge them together. However, the number of zero-effort impostors and spoofing attacks is highly dependent on the database and follows the database protocol. Hence, the ratio of the two classes into the super-class is different for different databases and can not be controlled. Furthermore, the super-class tends to be biased towards the component with more samples. For example, in a typical biometric verification database with N identities and M samples per identity, the number of zero-effort impostors will be $N \times (N-1) \times M$. On the other hand, if there is a single spoofing attack for any genuine sample in the database, the number of spoofing attacks will be $N \times M$. The correct ratio of zero-effort impostors and spoofing attacks into the super-class of negatives can not be known in advance. Any ratio of the two negative classes may be valid depending on the deployment conditions. For example, in highly supervised conditions, like airport control gates, spoofing attacks are more difficult to perform, and hence unlikely. On the other hand, unsupervised verification systems of portable devices are much more exposed to spoofing attacks. Thus, tuning the operating point of any system depends on its expected usage scenario.

The Expected Performance and Spoofability (EPS) framework is developed taking in consideration the above observations and is inspired by Expected Performance Curve (EPC) [1]. It introduces a parameter $\omega \in [0, 1]$, which balances between the error rates associated with the spoofing attacks (SFAR) and zero-effort impostors (FAR). The parameter ω can thus be interpreted as the cost, or the prior, of spoofing attacks relative to the cost, or prior, of zero-effort impostors, offering a solution to the problem of how to combine the classes of samples that need to be rejected.

Using the parameter ω, EPS framework introduces a metric called FAR_ω, which is a weighted error rate for the two negative classes (zero-effort impostors and spoofing attacks). It is calculated as in Eq. 1.

$$\mathrm{FAR}_{\omega} = \omega \cdot \mathrm{SFAR} + (1 - \omega) \cdot \mathrm{FAR} \tag{1}$$

The metric for assessing the global performance of the system takes into account another parameter, $\beta \in [0, 1]$, which balances between the error rates associated with the negative super class (FAR_{ω}) and the genuine users (FRR). It can be interpreted as the cost, or prior, of the negative super-class with respect to the positive class. Then, the metric for assessing the global performance is defined as Weighted Error Rate, $\mathrm{WER}_{\omega,\beta}$ and is calculated as in Eq. 2.

$$\mathrm{WER}_{\omega,\beta} = \beta \cdot \mathrm{FAR}_{\omega} + (1 - \beta) \cdot \mathrm{FRR} \tag{2}$$

A special case of $\mathrm{WER}_{\omega,\beta}$, obtained by assigning equal cost $\beta = 0.5$ to FAR_{w} and FRR can be defined as HTER_{ω} and computed as in Eq. 3.

$$\mathrm{HTER}_{\omega} = \frac{\mathrm{FAR}_{\omega} + \mathrm{FRR}}{2} \tag{3}$$

2.3 Determining the decision threshold

The last requirement is to define a criteria for determining the optimal classification threshold $\tau^*_{\omega,\beta}$. It depends on both parameters ω and β and is chosen to minimize the weighted difference between FAR_{ω} and FRR on the development set, as in Eq. 4.

$$\tau^*_{\omega,\beta} = \arg\min_{\tau} |\beta \cdot \mathrm{FAR}_{\omega}(\tau, \mathcal{D}_{dev}) - (1 - \beta) \cdot \mathrm{FRR}(\tau, \mathcal{D}_{dev})| \tag{4}$$

The decision threshold is a tunable parameter, and as such, it should be always determined using the development (validation) set of the database (denoted as $\mathcal{D}_{dev}$ in Eq. 4). On the other hand, usually the performance metric and error rates defined in Section 2.1 and Section 2.2 should be reported using the test set $\mathcal{D}_{test}$.

The parameter ω, as mentioned before, could be interpreted as relative cost of the error rate related to spoofing attacks. Alternatively, it could be connected to the expected relative number of spoofing attacks among all the negative samples presented to the system. In other words, it could be understood as the prior probability of the system being under a spoofing attack when it is misused. If it is expected that there is no danger of spoofing attacks for some particular setup, it can be set to 0. When it is expected that some portion of the illegitimate accesses to the system will be spoofing attacks, ω will reflect their prior and ensure they are not neglected in the process of determining the decision threshold.

Most importantly, the parameter ω ensures that the spoofing attacks are taken into consideration in the threshold decision process, with the prior associated to them. This is not true for a traditional evaluation of biometric recognition systems, where the threshold

is obtained using EER criteria based solely in FAR and FRR. However, this is of great importance, as ignoring the spoofing attacks in the threshold decision process makes the systems unnecessarily more vulnerable to spoofing, especially in the case when spoofing attacks are highly probable.

The parameter β could be interpreted as the relative cost of the error rate related to the negative class consisting of both zero-effort impostors and spoofing attacks. This parameter can be controlled according to the needs or to the deployment scenario of the system. For example, if we want to reduce the wrong acceptance of samples to the minimum, while allowing increased number of rejected genuine users, we need to penalize FAR_ω by setting β as close as possible to 1.

2.4 Visualization using Expected Performance and Spoofability Curve (EPSC)

The EPS framework can compute error rates for a range of decision thresholds obtained by varying the parameters ω and β. They can be plotted using Expected Performance and Spoofability Curve (EPSC), showing $\text{WER}_{\omega,\beta}$ with respect to one of the parameters, while the other parameter is fixed to a predefined value. For example, one can fix the parameter $\beta = \beta_0$ and draw a 2D curve which plots $\text{WER}_{\omega,\beta}$ on the ordinate with respect to the varying parameter ω on the abscissa. Having in mind that the relative cost given to FAR_ω and FRR depends mostly on the security preferences for the system, it is not difficult to imagine that particular values for β can be selected by an expert. Similarly, if the cost of SFAR and FAR or the prior of spoofing attacks with regards to the zero-effort impostors can be precisely estimated for a particular application, one can set $\omega = \omega_0$ and draw a 2D curve plotting $\text{WER}_{\omega,\beta}$ on the ordinate, with respect to the varying parameter β on the abscissa.

Besides $\text{WER}_{\omega,\beta}$, EPSC can present other error rates which are of interest. For example, plotting SFAR can show how the system's robustness to spoofing changes with regards to ω or β. Alternatively, to report on all the incorrectly accepted samples, FAR_ω can be plotted using EPSC.

Figure 2 and Figure 3 give an illustration of the EPSC plotting the error rates $\text{WER}_{\omega,\beta}$ and SFAR as function of the parameters ω and β, respectively. The plots are generated for a hypothetical verification system.

Figure 2a and Figure 2b show $\text{WER}_{\omega,\beta}$ and SFAR with respect to ω for three predefined values of β. The blue curve on Figure 2a, corresponding to $\beta = 0.5$, is equivalent to HTER_ω. The left-most points of the curves correspond to $\omega = 0$, meaning that the decision threshold is obtained disregarding the spoofing attacks as possible input. For the particular hypothetical system and all the three considered values of β, this point corresponds to low $\text{WER}_{\omega,\beta}$, which indicates a system with good verification capabilities, but very high SFAR due to the high overlap of the scores of spoofing attacks and genuine users.

Figure 2: EPSC of a hypothetical biometric verification system under spoofing attacks, parameterized over ω

(a) $\text{WER}_{\omega,\beta}$ (b) SFAR

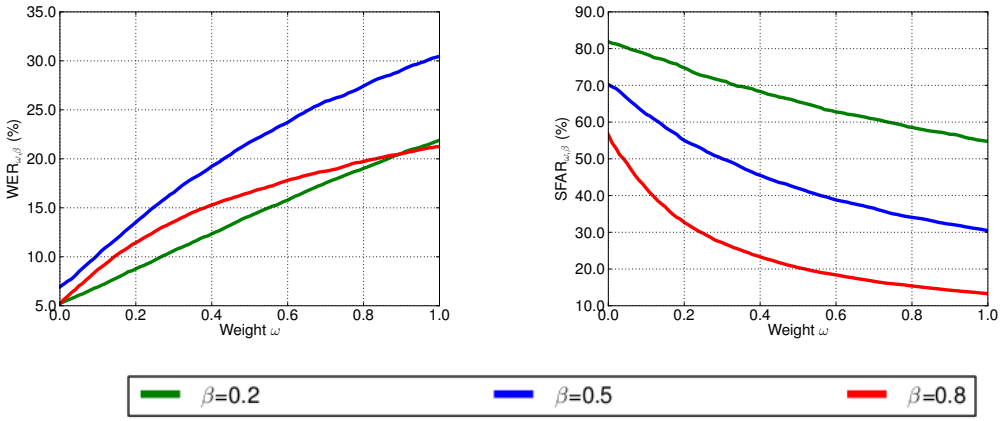
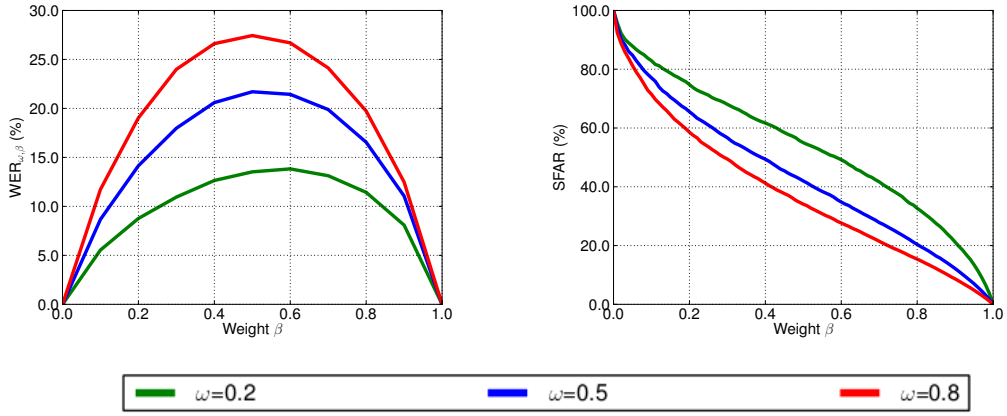


Figure 3: EPSC of a hypothetical biometric verification system under spoofing attacks, parameterized over β

(a) $\text{WER}_{\omega,\beta}$ (b) SFAR



As we increase ω, we give weight to the spoofing attacks so that they have a role in the threshold decision process. This decreases the number of spoofing attacks that pass the system, which explains why SFAR decreases with increasing ω. However, the additional caution for the danger of spoofing attacks unavoidably comes with the price of more rejected genuine users and thus higher $\text{WER}_{\omega,\beta}$. A system with high robustness to

spoofing attacks will show as mild increase of $\text{WER}_{\omega,\beta}$ as possible, with as steep decrease of SFAR as possible.

Figure 3a and Figure 3b show EPSC parameterized over the varying parameter β, for three predefined values of ω. For the extreme cases where $\beta = 0$ and $\beta = 1$, $\text{WER}_{\omega,\beta}$ is 0 because the threshold is determined to minimize the error rate solely associated with the positive or the negative class, respectively. In the case of $\beta = 0$, this results in a successful passing through of all the spoofing attacks.

Considering the spoofing attacks when calculating the decision threshold means taking additional precautions against them. As a result of this, the threshold obtained using EPS framework is better adapted to the input that is expected, contributing to systems with better performance and lower spoofing vulnerability, than systems whose decision threshold has been determined in different way.

The EPSC has the important property of unbiased system comparison, because it reports the error rates *a priori*. Since the threshold is always determined using the development set, and the error rates are reported using the test set, one can estimate the expected error rates and spoofability of the system in an unbiased way, on data which has not been seen before. The expected error rates can be reported for a particular value or range of values of the parameters ω and β which are of interest in a particular application. Moreover, EPSC allows for easy and unbiased comparison of verification systems with regards to their performance and robustness to spoofing, simply by comparing the EPSC for the two systems on the same plot. Even more, one can compare verification systems range-wise: which one performs better for a range of values of ω or β.

Finally, if a single number is needed to describe the performance of a system, Area Under EPSC (AUE) metric is defined, and can be computed for a fixed β or ω. For example, for a fixed β, it represents the average expected $\text{WER}_{\omega,\beta}$ for all values of ω and is computed using Eq. 5. The formula to compute AUE for fixed ω and varying β follows accordingly. Between two systems, better is the one which achieves smaller AUE.

$$\text{AUE} = \int_{\omega \in [0,1]} \text{WER}_{\omega,\beta}(\tau^*_{\omega,\beta}, \mathcal{D}_{test}) \mathrm{d}\omega \tag{5}$$

The AUE can be computed in between certain bounds $a, b \in [0, 1]; a < b$, enabling to compare two systems depending on the required range of the varying parameter.

2.5 Examples of EPSC: face mode

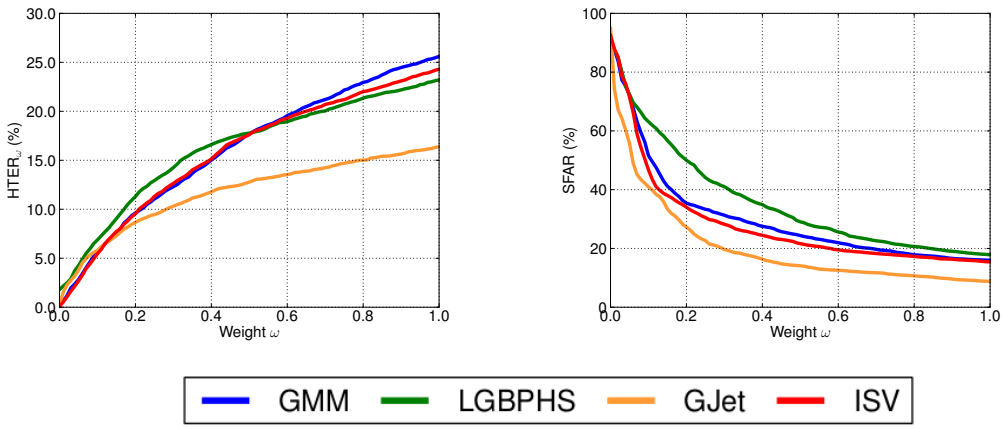
To exemplify the usage of the EPSC and demonstrate how to interpret its results, this section compares the performance and spoofing vulnerability of four state-of-the-art biometric recognition systems working in the face mode. Their performance is tested with the face spoofing database Replay-Attack [4].

The first one is a Gaussian Mixture Model (GMM) based system which extracts Discrete Cosine Transform (DCT) features from the input images [2]. The second one, called Local Gabor Binary Pattern Histogram Sequences (LGBPHS) [12], calculates Local Binary Patterns (LBP) histograms over the input images convoluted with Gabor wavelets, and computes the similarity scores using χ^2 measure. The third considered system is based on [11] and compares Gabor jets extracted from different positions and put into a single rectangular grid graph (GJet) [6]. Finally, DCT features are used once again in the fourth system, to create Universal Background Model and to estimate a linear subspace of the within-class variability [10]. We will refer to this system as Inter-Session Variability modeling (ISV).

For simplicity, the EPSC in the following figures are calculated for fixed $\beta = 0.5$ and they are parameterized by the parameter ω. The EPSC given in Figure 4, reports HTER_ω and SFAR for the four baseline systems, for a threshold which considers the relative probability of spoofing attacks, encoded in the parameter ω.

Figure 4: EPSC to compare baseline face recognition systems

(a) HTER_ω (b) SFAR



By plotting the EPSC for the four systems together, we can directly compare their overall performance (Figure 4a) and vulnerability to spoofing (Figure 4b). Even more, this comparison can be done for a range of values of ω. Depending on the prior of the spoofing attacks for a particular application, one can select the best system based on this comparison.

For example, comparing the HTER_ω values in Figure 4a, we observe that the ISV system is best performing in verification only as long as the spoofing attacks appear with a very small probability. After a certain value of ω, GJet shows the best verification performance. The same applies to the vulnerability to spoofing (Figure 4b): while being

the most vulnerable when $\omega \approx 0$, GJet displays the smallest values of SFAR for larger values of ω.

Similar conclusions about the overall performance of the systems can be taken by considering the AUE values in Table 1 as well. With AUE value lower than the other systems, GJet performs the best. However, note that this value is computed over the full range of ω. AUE values can be computed and compared for a specific range of ω as well.

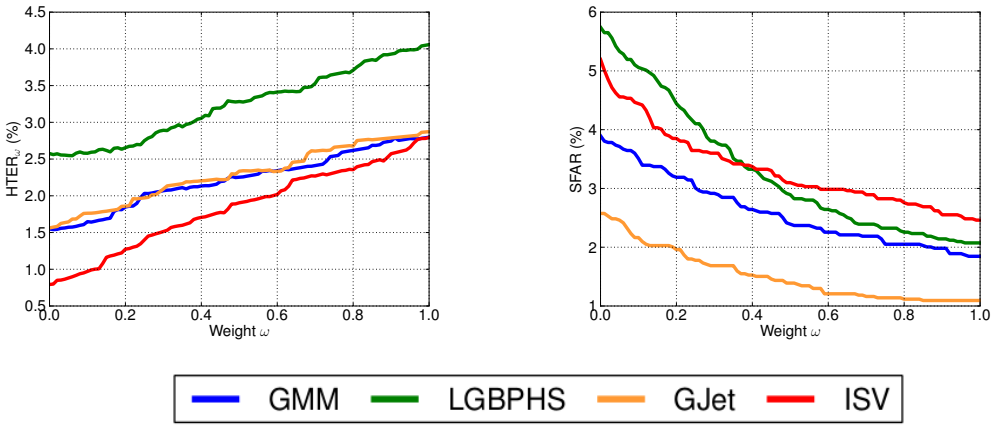
System	UBM	LGBPHS	GJet	ISV
AUE	0.161	0.162	0.117	0.158

Table 1: AUE values for baseline face recognition systems

Figure 4b shows that the baseline face recognition systems studied here by themselves show high vulnerability to spoofing. In the experiment that follows their robustness to spoofing is increased by fusing with several anti-spoofing systems using Polynomial Logistic Regression (PLR), as suggested in [5]. The EPSC of these systems after fusion are compared in Figure 5.

Figure 5: EPSC to compare face recognition systems fused with anti-spoofing systems

(a) HTER_ω (b) SFAR



The comparison between the EPSC for the baseline (Figure 4a) and the fused systems (Figure 5a), confirms that fusion is highly beneficial to the systems' robustness to spoofing.

The comparison of SFAR of the fused systems, given in Figure 5a, suggests that GJet baseline fused with the anti-spoofing systems is the least vulnerable to spoofing, regardless of ω. If we summarize both the verification performance and spoofability of the systems

into HTER_ω, Figure 5a suggests that ISV baseline fused with all the anti-spoofing systems performs the best.

By comparing the AUE values computed over the full range of ω given in Table 2, ISV shows superior performance among all the fused systems. Certainly, in a general case, this does not have to be true for a specific subrange of ω.

System	UBM	LGBPHS	GJet	ISV
AUE	0.022	0.032	0.023	0.018

Table 2: AUE values for face recognition systems fused with anti-spoofing systems

3 Metric for vulnerability Assessment

3.1 General approach

Within the current version of the Common Criteria, a number of aspects are used to characterize an attack. According to the original criteria in [3] the attack is rated after

- the elapsed **time** for the attack,
- the **expertise** needed to perform the attack,
- the **knowledge** of the TOE that is required,
- the **Window of Opportunity** that is needed,
- the **equipment** that is needed.

Each of the aspects is rated after a defined metric and ends up in a numeric value. The sum of those values stands for the difficulty of the overall attack. If a system resists all attacks of a certain value, the system fulfils a well defined level of resistance. Common Criteria further defines levels of resistances for a Target of Evaluation that are characterized in form of the same numeric values. With other words: In order to belong into a certain class or resistance a system will need to withstand all possible attacks of a certain value.

This approach for rating attacks is straightforward and often used. There is however an important drawback of this approach:

Many attacks comprise dedicated phases. It is often the case that extensive research activities are required in order to identify a potential attack path. Once such an attack path is known and has been published, the exploitation of such an attack may be very simple and fast. This is not just the case for spoofing attacks to biometric systems but can also be found in many other technologies.

This situation therefore lead to alternative rating schemes for attacks, specifically in the area of smart cards where the rating of an attack is explicitly structured after an identification phase and an exploitation phase (see [8]).

The identification phase corresponds to the effort required to create the attack, and to demonstrate that it can be successfully applied to the system (including setting up or building any necessary test equipment). The demonstration that the attack can be successfully applied needs to consider any difficulties in expanding a result shown in the laboratory to create a useful attack. One of the outputs from this phase could be a script that gives a step-by-step description of how to carry out the attack - this script is assumed to be used in the exploitation part.

The exploitation phase corresponds to achieving the attack on an instance of the system using the analysis and techniques defined in the identification part of an attack. Could be assumed that a different attacker carries out the exploitation, the technique (and relevant background information) could be available for exploitation in the form of a script or set of instructions defined during the identification of the attack. This type of script is assumed to identify the necessary equipment and, for example, mathematical techniques used in the analysis. For example, it is assumed that the script identifies such things as the physical point at which to apply a brute force attack, and hence in the exploitation phase the attacker does not have to spend significant time to find the correct point at which to apply the attack. Furthermore, this same information may also reduce the exploitation requirement to one of time measurement, whereas the identification phase may have required reverse engineering of hardware or software information from power data hence the expertise requirement may be reduced.

This paper proposes to utilize a similar approach for the rating of direct attacks against biometric devices. It introduces the details and a classification for the aforementioned characteristics of an attack. While the metric in this paper introduces all relevant details of the characteristics of an attack, it does not yet apply any concrete rating (in form of numeric values) in order to classify an attack. This part is left to deliverable D6.5

Along the lines of the approach from [8] the metric defined in this document describes the level of a spoofing attack under consideration of the following aspects

- The elapsed **time** for the attack,
- the **expertise** needed to perform or prepare the attack,
- the **knowledge** of the TOE that is required,
- the **Window of Opportunity** that is needed,
- the **equipment** that is needed.

during the identification and the exploitation phase. However, not all aspects are relevant for both phases of an attack. The following table summarizes which aspects are relevant for each of the both phases.

	Identification phase	Exploitation Phase
Time	X	-
Knowledge	X	
Window of Opportunity	X	X
Equipment	X	X

More details on the various aspects can be found in the following subchapters.

3.2 Elapsed Time

This factor is only applied to the Identification step. The time necessary to perform the attack on the final system (Exploitation) is taken into account in the Window of opportunity factor.

It corresponds to the time required to create the attack, and to demonstrate that it can be successfully applied to the biometric system (including setting up or building the concrete spoof to be used). The demonstration that the attack can be successfully applied needs to consider any difficulties in expanding a result shown in the laboratory to create a useful attack.

One of the outputs from identification could be, for instance, be a concrete spoof to be used with a certain biometric system along with a description on how to reproduce the spoof.

It may also include information on possible variations during the construction of the spoof that can be required in order to overcome potential countermeasure of the biometric system. More concretely, elapsed time in identification corresponds to the time spent to find the so called golden fake [1] and to define the method to build it from for example a fingerprint (with or without the collaboration of the user).

3.3 Elapsed Time

For attacks in which a user shall be impersonated by an attacker, the attacker will also have to acquire the biometric characteristic of the genuine user that can be copied to the spoof. The time and effort that is needed for this acquisition also counts for this rating. While acquiring the biometric characteristic may be very simple for some biometric characteristic (e.g. for a fingerprint that can usually be acquired from latent fingerprints left somewhere) it may require more effort for other characteristics.

The identification phase also includes the time needed for the analysis of any countermeasures that a system may provides in order to resist spoofed biometric characteristics.

3.4 Expertise

This factor refers to the level of proficiency required by the attacker and the general knowledge that they possess, not specific of the system being attacked. A suggested rating for this metric is:

[1] In the case of spoofing attacks, for example with gummy fingers, we will consider the existence of a golden fake, manufactured with a specific material, which, once identified (in the identification phase), is able to break a given scanner/system with very few attempts for almost all the cases.

Layman no real expertise needed, any person with a regular level of education is capable of performing or preparing the attack.

Proficient some advanced knowledge in certain specific topics (e.g., use of specific tools) is required. Good knowledge of the state-of-the-art of attacks. The person is capable of adapting known attacks to their needs.

Expert a specific preparation in multiple areas such as pattern recognition, computer vision or optimization is needed in order to carry out the attack. The person is capable of generating their own new attacking algorithms.

Multiple experts the attack needs the collaboration of several people with high level expertise in different fields (e.g., material sciences, cryptanalysis, physics, etc.)

The factor of expertise may have a strong correlation with the time that is needed to prepare an attack. A well experienced attacker may be way faster in identifying and building a spoof that works with a certain biometric system than a layman. This also leads to the situation that multiple ways of identifying an attack may be worth to be considered and compared. Staying in this specific scenario it can be worth to compare the rating of a layman who needs a significant amount of time to prepare an attack. Please refer to 3.8 for more information.

3.5 Knowledge of the TOE

This factor refers to the amount of knowledge about the attacked system required to identify and/or perform the attack.

For instance, the format of the acquired samples, size and resolution of acquisition systems, specific format of templates, but also specifications and implementation of countermeasures are knowledge that could be required to set up an attack.

This information could be publicly available at the website of the sensor manufacturer or protected (distributed to stakeholders under NDA or even classified inside the company).

Ratings are:

Public information which is fairly easy to obtain (e.g., in the internet).

Restricted information which is only shared by the developer and organizations which are using the system, usually under a non-disclosure agreement.

Confidential information which is only available within the organization that develops the system and is in no case shared outside it.

Critical it refers to information which is only available to certain people or groups within the organization which develops the system.

Special attention should be paid to possible countermeasures that may be implemented in the system and whether it is necessary or not to have knowledge of their existence in order to be successful in a given attack. This aspect may have a strong correlation to the factor of elapsed time. If an attacker is not aware of the concrete countermeasures implemented by a system it may take significantly more time to prepare a successful attack then when knowing the detailed specification.

It is assumed that all the knowledge required to perform the attack is gained during the identification phase and prepared for the exploitation. Therefore, this factor is not applied to the exploitation phase.

3.6 Window of Opportunity

This factor refers to the difficulty to access the system either to prepare the attack (in identification phase) or to perform it on the target system.

For the identification phase, elements that should be taken into account include the easiness to buy the same equipment (with and without countermeasures) that should be attacked.

For exploitation phase, both technical (such known/unknown tuning) and organizational measures (presence of a guard, ability to physically modify the target, limited number of tries, etc.) should be taken into account.

This factor is not just expressed in terms of time. Proposed values are:

Easy There is no strong constraint for the attacker to buy the system (at a reasonable price) to prepare its attack (identification phase). For the exploitation phase, only a few (1 to 3 for example) tries are required and the spoof will work with a sufficient reliability.

Medium For the identification phase, specialized distribution schemes exist (not available to individuals). For exploitation phase, either a tuning of the attack for the final system is required (unknown parametrization of countermeasures for example), or several tries (5-10) are necessary to reach a good level of success.

Difficult For the identification phase, the system is not available except for identified users. For exploitation, for example spoofs must be adapted to the (unknown) specific tuning, or there is only a very limited chance for the spoof to be accepted (more than 50 tries requested), or the system need physical modification (for example physically accessing a hidden signal significant of the matching rate).

3.7 Equipment

This factor refers to the type of equipment required to perform the attack. A suggested rating is:

Standard equipment which is affordable, easy to obtain and simple to operate (e.g., computer, video cameras or mobile phones).

Specialized this refers to fairly expensive equipment, not available in standard markets and which requires of some specific formation to be used (e.g. a spectrum analyser).

Bespoke this refer to very expensive equipment with controlled access; for example, research printing systems with specific ink definition and flexible support adaptation. In addition, if more than one specialized equipment is required to perform different parts of the attack, this value should be used.

3.8 Correlations between rating aspects

It is essential to understand that there is not just one way to rate a specific attack. Rating a specific attack does not presuppose to have a dedicated attacker in mind. An attack usually has a certain and well defined target. Such a target could for example be to impersonate a genuine user.

In order to achieve this target an attacker has many different options and ways and one should assume that there is more than one attacker existing.

Even if a very concrete scenario is observed (e.g. spoofing a system by the use of a gummy finger) the attack can be rated in many different ways. One could for example start with an attacker who is a layman and will need a significant amount of time to prepare a finger that would work for a certain system. On the other hand one can also start with a Proficient attacker who will only need a day to prepare such a spoof.

In such situations it is essential for the use of the metric as proposed in this document to rate **all possible and useful combinations** that an attacker could use in order to achieve their target. If multiple scenarios lead to different values (which is likely) the **smallest value should count** to characterize the attack as an attacker will always try to use the path of least resistance.

References

[1] S. Bengio, J. Mariéthoz, and M. Keller. The expected performance curve. In *International Conference on Machine Learning, ICML, Workshop on ROC Analysis in Machine Learning*, 2005.

[2] F. Cardinaux, C. Sanderson, and S. Marcel. Comparison of mlp and gmm classifiers for face verification on xm2vts. In *Proceedings of the 4th International Conference on AVBPA*, University of Surrey, Guildford, UK, 2003.

[3] CCBD, editor. *Common Criteria for Information Technology Security Evaluation*. 2009.

[4] I. Chingovska, A. Anjos, and S. Marcel. On the effectiveness of local binary patterns in face anti-spoofing. In *Proceedings of the 11th International Conference of the Biometrics Special Interes Group*, 2012.

[5] I. Chingovska, A. Anjos, and S. Marcel. Anti-spoofing in action: joint operation with a verification system. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Workshop on Biometrics*, June 2013.

[6] M. Günther, D. Haufe, and R. P. Würtz. Face recognition with disparity corrected Gabor phase differences. In *Artificial Neural Networks and Machine Learning*, volume 7552 of *Lecture Notes in Computer Science*, pages 411–418. Springer Berlin, 2012.

[7] P. A. Johnson, B. Tan, and S. Schuckers. Multimodal fusion vulnerability to non-zero (spoof) imposters. In *IEEE International Workshop on Information Forensics and Security*, 2010.

[8] J. I. Library. Application of attack potential to smartcards. 2.7, 2009.

[9] C. C. N. Technical Editor: National Cryptologic Centre (CCN. Characterizing attacks to fingerprint verification mechanisms. 3, 2011.

[10] R. Wallace, M. McLaren, C. McCool, and S. Marcel. Inter-session variability modelling and joint factor analysis for face authentication. In *International Joint Conference on Biometrics*, 2011.

[11] L. Wiskott, J.-M. Fellous, N. Krger, and C. V. D. Malsburg. Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis And Machine Inteligence*, 19:775–779, 1997.

[12] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang. Local gabor binary pattern histogram sequence (lgbphs): A novel non-statistical model for face representation and recognition. In *Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1 - Volume 01*, ICCV '05, pages 786–791. IEEE Computer Society, 2005.