Analyzing Handwriting Biometrics in Metadata Context

Tobias Scheidat, Franziska Wolf, Claus Vielhauer
Dept. of Computer Science, Univ. of Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany

ABSTRACT

In this article, methods for user recognition by online handwriting are experimentally analyzed using a combination of demographic data of users in relation to their handwriting habits. Online handwriting as a biometric method is characterized by having high variations of characteristics that influences the reliance and security of this method. These variations have not been researched in detail so far. Especially in cross-cultural application it is urgent to reveal the impact of personal background to security aspects in biometrics. Metadata represent the background of writers, by introducing cultural, biological and conditional (changing) aspects like fist language, country of origin, gender, handedness, experiences the influence handwriting and language skills. The goal is the revelation of intercultural impacts on handwriting in order to achieve higher security in biometrical systems. In our experiments, in order to achieve a relatively high coverage, 48 different handwriting tasks have been accomplished by 47 users from three countries (Germany, India and Italy) have been investigated with respect to the relations of metadata and biometric recognition performance. For this purpose, hypotheses have been formulated and have been evaluated using the measurement of well-known recognition error rates from biometrics. The evaluation addressed both: system reliance and security threads by skilled forgeries. For the later purpose, a novel forgery type is introduced, which applies the personal metadata to security aspects and includes new methods of security tests. Finally in our paper, we formulate recommendations for specific user groups and handwriting samples.

Keywords: biometrics, cross-cultural, metadata, online handwriting, skilled attacks, soft biometrics, verification

1. INTRODUCTION

Active methods in biometrics like voice and handwriting have achieved a new powerful position in user authentication systems recently. Previous researches already introduced to new methods of analyzing active biometric methods (voice and handwriting) in context of metadata. For example in handwriting recognition in [1] text is analyzed by using global image features and image indexing by automatic page analysis and segmentation. Image based on image-to-image similarity measure and text based on text-to-image score are used for retrieval of data. Metadata that describes both, technical and personal characteristics in order to relate both facts to the context of speech-based user authentication was introduced in [2]. Methods of analyzing handwritten documents regarding aspects of person related data like gender or ethic background have been made. In [3] handwriting of Indian users was analyzed for identification performance and quality based on demographic information about the writer. Different characters were ranked and the individual performance of characters for group identification was measured. Based on demographic data of gender an age the accumulated performance of characters could be achieved between 65% and 85%. Nevertheless only static type faces of handwriting have been researched so far. In our research as shown in recent work [4] we refer to dynamic handwriting data and relate this to metadata in order to investigate cultural impacts to enhance authentication security. Analysis on online handwriting data metadata can be used to enhance the evaluation of handwriting in intercultural context.

We now also introduce a new area of security application with respect to forgery levels. Metadata include contextual information regarding the experimental environment, as well as for the linguistic, cultural, ethic and educational background of a person, which is attributed to the biometric data. These new aspects of evaluating voice and handwriting considering cultural impacts like language, script and nationality are of special interest to technology. For example, different cultural background may affect usability and security (i.e. recognition accuracy) of a biometric system in international cross-cultural and cross-lingual context. In our previous work, we have presented a novel framework consisting of a metadata model and an acquisition methodology for an experimental environment based on the biometric modalities of speech and voice [5]. For the domain of face recognition, Phillips et al. describe in [6] a procedure and results in the Face Recognition Vendor Test 2002. For the tests a very large data set of 37,437 individuals was used and the data were examined also from so-called demographic aspects. For the best systems the identification rates for males were 6% to 9% points higher than that of females. Another result of the tests was that identification of older people is

Security, Steganography, and Watermarking of Multimedia Contents VIII, edited by Edward J. Delp III, Ping Wah Wong,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6072, 60720H, © 2006 SPIE-IS&T · 0277-786X/06/$15

easier than of younger people. The identification performance increases approximately 5% points for every ten years increase of age. These outcomes motivate further examinations in the area of metadata (here gender and age) also for other biometric traits like online handwriting.

New aspects of metadata based security analysis with respect to cultural and biological and new conditional metadata can be formulated. The main interest in both areas of research for us is the definition of user groups based on their metadata (e.g. by gender or native language). These groups are analyzed and compared by aspects of handwriting classified into relation of contend (here called semantic) and relation of data parameters (here called syntax) . The goal is to formulate metadata based recommendations for usability and security, which may enhance active biometric methods in future. The intercultural data collection is gained from the research project CultureTech ([7]), having project partners in Germany, Italy and India. Standardized handwriting samples and metadata have been collected in order to have a decent base for an analysis with focus on intercultural matters.

By analyzing both, user group's metadata spreading and their handwriting parameters, hypotheses are formulated that relate the metadata by intercultural aspects to the handwriting specifics. Evaluations can be done by referring to a well-established recognition measurement in biometrics, equal error rates (EER), with respect to inner and outer safety. Here, **inner safety** refers to verification and random attacks declares the system's security without the thread of forgery, whereas **outer safety** refers to varying attacks based on a hierarchy of sophistication. Even though the subject quantity in our evaluations cannot claim statistical completeness, bias estimations can be gained using these results and differing commendations concerning using samples can be formulated in order to achieve optimal security levels for each subject group due to metadata.

Figure 1 shows the general process of authentication based on enrollment data and test data. First, the enrollment data, consisting of handwriting data and metadata, is collected, in order to register a person in the system. Preprocessing and feature extraction leads to storage in a data base that organizes the collection. To investigate the impact of forgery to the authentication process, the inner and outer security is calculated using the handwriting data of genuine users and specially trained forgers. During the authentication process the handwriting data, of users and forgers, is matched with the enrollment data of the database after preprocessing and feature extraction. The matching uses the statistical method of the biometric hash algorithm described in [8] and leads to a decision that defines the system's outer and inner security aspects.

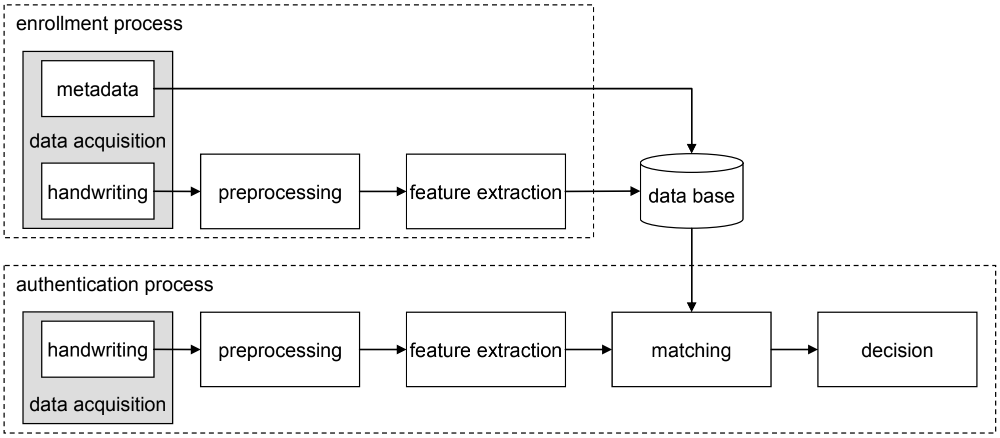

Figure 1: Function of enrollment and authentication process

In this work we will introduce in the terms of metadata. The classes of technical and non-technical, the biological and cultural, metadata are presented. Furthermore, the new class of conditional metadata is defined. Then, the handwriting samples giving base for the analyses are defined and the different tasks of given, individual and creative handwriting tasks are described. The experimental setup including error rates measurement and different forgery classes with focus on the new metadata regarding forgery type **blind meta attack** are defined. Putting together metadata analysis and handwriting data analysis of the intercultural working group CultureTech, three hypotheses suspecting intercultural handwriting aspects are given. Evaluations of these hypotheses with respect to inner and outer security lead to our distinct recommendations of best usable handwriting samples for differing user groups. Future work and possibly applications conclude this paper.

2. METADATA IN BIOMETRICAL CONTEXT

For analyzing the paradigms of dynamic handwriting with respect to intercultural aspects of the writer, both writer and handwriting data have to be classified and analyzed specifically. Metadata have been collected along the handwriting collection. For the recording of the online-handwriting data, a TabletPC (Toshiba Portégé) or a graphical tablet with integrated display (Wacom Cintiq15) has been used, collecting parameters like pen position and pressure, in order to determine statistical values like velocity, path lengths of writing, writing time, just to name a few. In the following we present these theoretical fundamentals of the classification and analysis of metadata and handwriting samples.

2.1 Classification of metadata

Metadata describe the surrounding of a biometrical data acquisition. As shown in figure 2, the metadata can be divided in two main classes: the technical and the non-technical metadata. In this work we focus on the non-technical aspects that describe the owner of biometrical data. In order to reach many aspects of a person, there are three sub-classes of non-technical metadata: the biological, the cultural and the conditional metadata.

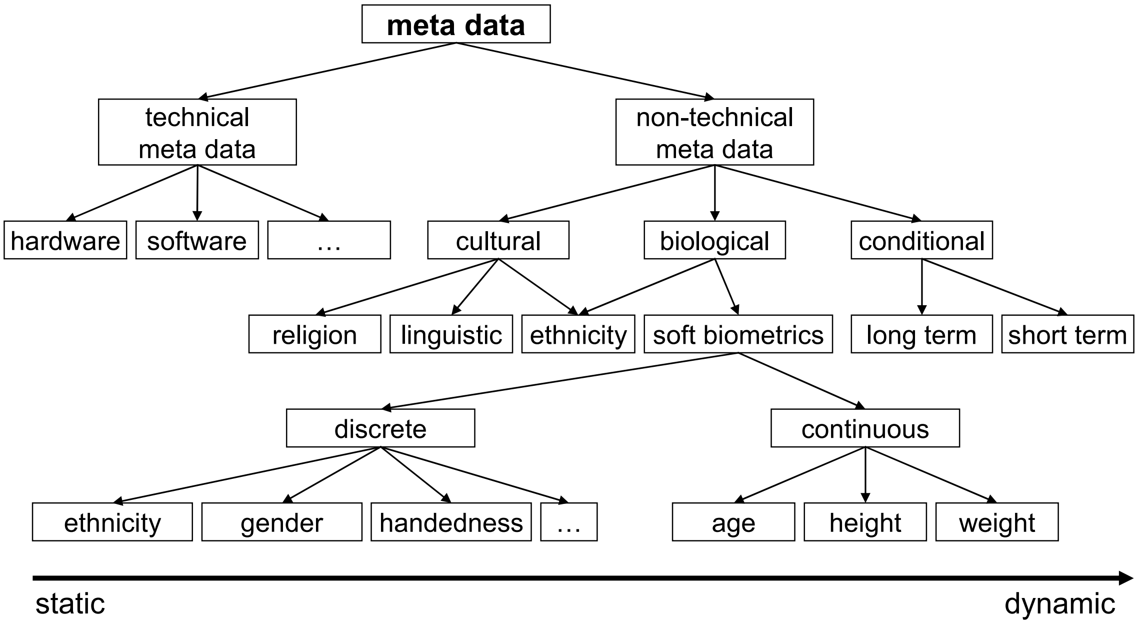

Figure 2: Classification of metadata

The biological metadata describes the sometimes called *soft biometrics* ([9]). Here, aspects that refer to the human body are mentioned, i.e. handedness, especially interesting in the context of handwriting, the age and the gender are collected using a software interface. These biological metadata can be used to research the physical impacts on handwriting data.

The cultural aspects collect the background of a person's development. Here aspects of provenance are as collected in form of the first and second languages and scripts known by the subjects. Languages and scripts can differ and therefore establishing a basis for differing handwriting as well.

Beyond these biological and cultural aspects, a new class of metadata could be established: the **conditional metadata**. This new class refers to the experiences and changing attitudes a person can have during its live, such as the actual physical state of a person during the data collection, that are classified as short term conditional metadata. It is known that states of fatigue or illness can affect the appearance of handwriting. Hence documentation about these states is urgent. Furthermore language changing experiences like stays abroad could affect handwriting as well and therefore should be investigated as long term conditional metadata. General attitudes towards biometrical usage could indicate reluctance to handwriting matters. The familiarity concerning the usage of computers or biometrical date can be related to handwriting habits as well. Table 1 gives an overview about the collected metadata in our work. In the following, we will focus on the biological metadata of gender, the cultural metadata of country of birth and the long term conditional metadata of stays abroad.

Having both, static biological and cultural metadata and changing conditional metadata a person can be characterized more detailed in order to achieve sophisticated recommendations for handwriting usage for different user groups. Provided that these metadata are sufficiently distributed the user groups can be compared to each other and therefore user group's specifics and handwriting can be led to intercultural aspects and general recommendations.

Table 1: Classification of metadata

Description of metadata	Metadata class
gender (female or male)	**non-technical, biological**
age	non-technical, biological
handedness (left or right)	non-technical, biological
ethnical background	non-technical, biological
religion	non-technical, cultural
education	non-technical, conditional, cultural
country of birth (ISO-3166),	**non-technical, cultural**
country of birth of parents (ISO-3166),	non-technical, cultural
countries of education (ISO-3166)	non-technical, cultural
first language (ISO-639)	non-technical, cultural
other languages (ISO-639)	non-technical, conditional, cultural
script of first language (ISO-15924)	non-technical, cultural
other scripts (ISO-15924)	non-technical, conditional, cultural
recorder (digitizer tablet, microphone)	technical
environment	technical
sample types	technical
language of test	non-technical, conditional
capital or cursive writing	non-technical, conditional, short term
data and time of recording	technical
experiences with computers / biometrical systems	non-technical, conditional, long term
stays abroad	**non-technical, conditional, long term**
attitudes towards biometrical systems	non-technical, conditional, long term
Physical state during data acquisition	non-technical, conditional, short term
States of fatigue, concentration and distraction during data acquisition	non-technical, conditional, short term

2.2 Handwriting data collection and analysis

Before the relation of personal data to handwriting can be done, a decent analysis of the handwriting data itself has to be done. Handwriting data was collected working together in an intercultural researching group of the CultureTech project [7]. The volunteers in Germany, Italy and India followed a guideline that structured the English tasks. Additionally, the German group recorded the same tasks in German in order to enable inter-language comparisons. These 48 tasks can be grouped in individual, creative and given tasks and are classified in the following table 2.

In order to analyze the collected handwriting tasks regarding their classification, two ways of analysis have been developed: the syntactical and the semantical analysis.

The **syntactical analysis** compares the handwriting parameters, which are recorded by the digitizer tablet. The syntactical aspects of the handwriting data determine the look and the dynamic of the handwriting and therefore the general security usage of handwriting in general. On this syntactical data the authentication and decision process is based using the biometric hash algorithm as introduced by Vielhauer et all. in [8] for the first time and enhanced in [10] and [11]. This algorithm extracts the online handwriting data into 68 different statistical features, like path-length, time of writing or pressure. By comparing these 68 parameters, varieties of handwriting and therefore the users can be measured.

Table 2: Classification of handwriting tasks

Number of task	Task	Classification of task
1	Signature	Individual
2	Alias	Individual
4	Pass phrase	Individual, creative
5	Symbol	Individual, creative
6	Numbers 0-9	Given
7	Latin Alphabet	Given
8 – 12	Questions to answer	Individual
13 – 15; 19 – 34	Words	Given
3; 35 – 41	Numbers	Given
16 – 18; 42 – 48	Sentences	Given

Nevertheless the syntactical aspects do not consider the classification of tasks that are given above in table 2. The whole content of the writing can not be comprehended by using these parameters exclusively. Therefore a subjective measurement has been accomplished in order to include the semantic matters as well. The **semantical analysis** researches handwriting samples by their semantic regarding following aspects: Individual and creative samples have been checked for their meanings and special aspects. For example if abbreviations have been used for signatures, what kind of symbols has been used and what has been chosen as password. At the individual questions the given answers are compared. All the samples are analyzed by their visual appearance and matters of completeness.

Having the analysis of user groups and their associate handwriting aspects, connections are examined. By formulating hypotheses that attend to security matters of handwriting specifics of intercultural user groups these connections are fulfilled. Examples are given in chapter 4.

3. EXPERIMENTAL SETUP

This chapter gives an overview of database and methodology on which the tests are based. The examined handwriting semantics are introduced and error rates used for the determination of the performance are explained. In addition, the test procedure of producing enrollments, verifications and forgeries are described.

3.1 Test database

The test plan allows the capturing of two different languages (German, English) in order to provide the possibility of cross-lingual comparison. Audio samples have been recorded also, following the same structure then the handwriting plan on which we will focus in detail. In total, each test person had to write 48 different semantics with ten iterations each. The semantics consisting of the signature, arbitrary pass phrase and symbol, personal questions and words, sentences and numbers (see table 2 in chapter 2.2). 28 German, 11 Italian and 8 Indian test persons participated in our experiments and samples have been recorded in German (28 Germans) and English (11 Germans, 11 Italians, 8 Indians). All handwriting samples have been captured either on a Toshiba TabletPC or a Wacom Cintiq15 tablet. These devices were selected since they have an active display. The advantage of an active display is that the handwriting trace is shown on the computer display at the pen-time location and in near real-time, a concept which is also sometimes called **digital**

ink. Through this technology, the quality of both the enrollments and the verifications can be considered relatively high. The quality of forgeries additionally also improve by this technique, since dealing with the active displays corresponds to the natural human writing behavior.

Metadata of all classes (biological, cultural and conditional) could be collected from the German and Italian volunteers. From the Indian test subjects only the biological and cultural metadata are available. In total, 28,170 Samples of handwriting and 7,770 samples of voice could be collected in the described manner and used for the analysis. The audio samples shall be used in upcoming work regarding the metadata and in fusion with handwriting. For evaluating the data efficiently the amount of analyzed tasks had to be limited for the first experiments. Thus individual, creative and given tasks had to be selected to following representative excerpts:

- the signature as the classical handwriting for authentication (individual task),
- the arbitrary symbol as an creative non-verbal sample (creative task),
- two numbers (*8710*, *77993*) and two words (*bird*, *communication*) in different length (given tasks) and
- one sentence (*Where are you going?*) as a combination of several words (given task).

3.2 Test methodology

For analyzing online handwriting with respect to verifications and forgeries, Zoebisch and Vielhauer suggested a test methodology in [12] which we have adopted. Enrollments, verifications and forgeries are captured to be able to estimate the authentication performance of an online handwriting system. During the enrollment process, the data of users are registered in the system for the first time for later use as reference data. Verification samples are used for the simulation of the authentication users by the system. The forgeries are subdivided into four groups depending on the strength of the attack. This strength is based on the attacker's knowledge about the original handwriting samples. The **random attack** is based on all verification data except the data of each examined user. In the **blind attack** scenario, the forger only knows which handwriting semantic (e.g. signature) he or she shall falsify. If the forger is in possession of a blue-print (offline representation) of the original handwriting and can trace the shape during the writing process, the forgery is called a **low force attack**. In addition to this, in case the forger has the complete available information about the original sample at the **brute force attack**. Such information can for example be the temporal characteristics of position or pressure. In addition to the mentioned attacks we have developed a new attack type based on the metadata, the **blind meta attack**. Based on the level of given information, it can be ranked between blind attack and low force attack strengths. The forgers here receive additional background information about the persons to be falsified. This information can be the gender or grade of education, for example. The references are compared with the verifications and forgeries in order to determine the error rates of the system.

For our examinations we subdivided the test persons into groups based on the metadata. Based on these groups we have created hypotheses regarding the way of writing. We then have tried to verify the hypotheses by the determination of biometrical error rates. The **false non match rate** (FNMR) indicates, how frequently authentic persons are rejected from the system. Basis for the calculation of FNMR is the comparison of the enrollment and verification data. The acceptance rate of non-authentic subjects is represented by the **false match rate** (FMR). The determination of the FMR is based on the relation of enrollment and forgery data of the different levels. In order to compare the results of different tests we used the **equal error rate** (EER). EER denotes the point in the error characteristics, where FNMR and FMR yield identical values. It needs to be stated however, that the EER do not represent the optimal operating point of a biometric system. The optimal operating point depends on the desired level of the security and/or comfort of the planned biometric system. The degree of correctness of the hypotheses can be determined by the comparison of the EERs of the semantics of the single groups. The indicated test data do not have a sufficient statistical significance. However, they shall demonstrate our fundamental procedure and motivate further extensive tests.

4. HYPOTHESES EVALUATION AND DESIGN RECOMMENDATIONS

According to the syntactical and semantical analysis described in chapter 2.2, three hypotheses were created by the authors. These shall be described and verified in this section. The hypotheses form only a first subjective choice of the authors and should be enhanced by further examinations in later work.

4.1 Analysis of metadata and handwriting

The following analysis does not show the statistical significance to give their recommendations a firm basis and the conclusions and recommendations are subjective implications of the researchers of the project. These analyses are subjective and manually done by the researchers and can hardly be considered as exhaustive. Nevertheless, we assume that a tendency of security aspects can be formulated to show new paths of analysis in this new formed area of biometrics and authentication systems. The analysis revealed three main scenarios using the metadata of gender, nationality and experience of stays abroad, to divide the whole user groups in three sub-groups, as shown in table 3.

The semantical and syntactical exploration showed some distinctive features that could be connected to these sub groups. As presented in section 3.2, the equal error rate of the biometric hash system is used to evaluate the hypotheses concerning the different handwriting aspects of the sub user groups with regard to security aspects of inner system reliance and forgery threads.

Table 3: Metadata used for building sub groups

Metadata for building sub groups	Sub-Groups	Number of volunteers
Nationality	All German volunteers	28
	All Italian volunteers	11
	All Indian volunteers	8
Gender	Female volunteers	14
	Male volunteers	33
Experience of stay abroad	Stays abroad experience (German only)	6
	No Stays abroad experience (German only)	22

4.2 Formulating hypotheses

The three hypotheses of the sub user groups regarding to nationality, gender and experiences abroad are shown next. As described before, two steps of evaluation have to be done for every hypotheses: firstly the evaluation of **inner security**, using the EER of verification (FNMR) and random (FMR) and secondly the evaluation of **outer security**, using the EER of verification (FNMR) and the four forgeries (FMR) that are shown in tables at every evaluation of hypotheses. For having equivalent basic values the number of test subjects in the compared sub groups was adjusted at an average of 7 to 14 test subjects each.

1st hypothesis

The first hypothesis refers to different cultural background of the sub groups. The analysis of semantic and syntax showed that signatures and numbers could be divided by the metadata of nationality. Signatures showed a great amount of difference, like Indian volunteers writing in non-Latin letters contrary to European users. Furthermore, the Indian test subjects wrote numbers in a more reduced style than all of the Italian and some of the German volunteers did. The first hypotheses can thus be formulated as follows:

Cultural differences can be shown by comparing; therefore differing recommendations can be made for each sub group.

For evaluation of the hypotheses the handwriting samples of the three nationalities (G=German, It=Italian, In=Indian) are used to calculate the error rates of the system and the forgery threads.

Table 4 shows the EERs for the national sub groups, the entries of special interest are marked in the table for matter of clarity. The first column (EER Random) shows the values for the EERs of the groups and their Random test of verification, the ongoing columns show the rates of the attacks. In the rows EER is represented into dependence to the semantics. The inner security aspects showed differences of the national groups could be made. In general it was evident that Indian handwriting data was more reliable to the system than the European handwriting data, what is especially true for the Random test of verification at the signature (EER(GE)=0.04, EER(IT)=0.10, EER(IND)=0.01) and the number *8710* (EER(GE)=0.19, EER(IT)=0.15, EER(IND)=0.10).

Table 4: EERs of all tasks for national sub groups

	EER Random			EER Blind			EER Meta			EER Low Force			EER Brute Force		
Semantic	**GE**	**IT**	**IND**	**GE**	**IT**	**IND**	**GE**	**IT**	**IND**	**GE**	**IT**	**IND**	**GE**	**IT**	**IND**
Signature	**0.04**	**0.10**	**0.01**	0.06	0.07	0.04	**0.06**	**0.26**	**0.08**	0.18	0.26	0.04	0.22	0.25	0.09
Symbol	0.07	0.08	0.09	0.07	0.05	0.06	0.12	0.07	0.12	0.20	0.18	0.15	0.22	0.22	0.15
8710	**0.19**	**0.15**	**0.10**	0.06	0.12	0.05	0.07	0.20	0.12	0.07	0.16	0.09	0.07	0.12	0.09
77993	0.10	0.22	0.06	**0.12**	**0.01**	**0.02**	**0.10**	**0.01**	**0.04**	**0.14**	**0.03**	**0.07**	**0.16**	**0.08**	**0.10**
Bird	0.05	0.10	0.10	0.16	0.07	0.06	0.09	0.10	0.07	0.13	0.14	0.07	0.20	0.80	0.04
Communication	0.10	0.10	0.14	0.06	0.06	0.02	0.14	0.03	0.03	0.10	0.07	0.04	0.08	0.07	0.07
Where are you going?	0.09	0.10	0.05	0.09	0.06	0.02	0.06	0.03	0.01	0.14	0.08	0.03	0.10	0.06	0.03

The evaluations referring to the outer safety are now described. For the sake of clarity only some significant examples of the forgery's EER are discussed here. The Forgeries of numbers show some interesting features. The attacks to the Italian handwriting samples show that for Italian users, the meta attack on the signature is the most effective, but the same attack on the other Italian handwriting samples hardly shows any success. Indian and German writers are more endangered by giving metadata data to forgers. All of the national groups show the best results by using the sentences, i.e. these semantics appear to be the most secure usage facing meta forgery. Consequently, recommendations for general usage can be formulated not only for the blind meta attack, but for all inner and outer security pre-conditions.

Table 5: Recommendations for national sub groups

Sub groups	**Random**	**Blind**	**Meta**	**Low Force**	**Brute Force**
German	Signature	Middle word	Sentence	Short number	Short number
Italian	Symbol	Symbol	Sentence	Long number	Short word/ Sentence
Indian	Signature	Sentence	Sentence	Sentence	Sentence

Differences of best and worst usage can be revealed, and therefore recommendations given for every national group. Comparing all handwriting samples of all sub groups and verification and forgery levels lead to formulation of the recommendations presented in table 5.

2nd hypothesis

The second hypothesis refers to the sub group build up only by the metadata of gender. Volunteers, not regarding their nationality, were classified in two groups: *female* and *male*. Comparing their semantical and syntactical features, differences showed especially at creative and individual samples. Therefore the second hypotheses could be formulated:

Distinct differences are detectable between handwriting samples of female and male writers. This is especially true for individual and short samples.

Inner security evaluation showed that the system reliance represented by the EER of verification and random test only differs in small amounts between the female and male user group. Nevertheless the male sub group showed slightly better results than the female group did. This was especially true for the individual and creative tasks. The women showed the best results at short words. Table 6 shows that there are primarily differences at the arbitrary symbol for females and males. The inner security test results show that, at the symbols, the EER of the male test subjects is lower (EER(symbol)=0.03) than with the female group (EER(symbol)=0.08). An also big difference in the EER was ascertained at the short word (Bird). The error rate of the female participants (EER(bird)=0.03) is considerably lower than at the male one (EER(bird)=0.13). The results of the other semantics differ on the inner security issues from each other among the two groups around a maximum value of 0.02.

Table 6: EERs for gender based sub groups

	EER Random		EER Blind		EER Meta		EER Low Force		EER Brute Force	
Semantic	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
Signature	0.03	0.04	**0.04**	**0.14**	**0.08**	**0.21**	**0.05**	**0.42**	**0.09**	**0.20**
Symbol	**0.03**	**0.08**	**0.03**	**0.10**	**0.11**	**0.15**	**0.17**	**0.19**	**0.21**	**0.19**
8710	0.10	0.12	0.05	0.14	0.10	0.12	0.08	0.18	0.08	0.18
77993	0.05	0.07	0.06	0.10	0.05	0.08	0.08	0.15	0.10	0.17
Bird	**0.13**	**0.03**	0.11	0.06	0.10	0.04	0.11	0.12	0.12	0.04
Communication	0.04	0.04	0.03	0.05	0.05	0.01	0.05	0.08	0.05	0.13
Where are you going?	0.08	0.09	0.06	0.10	0.03	0.03	0.18	0.11	0.15	0.12

The outer security survey is concentrated on the creative and individual tasks here. Comparing the EERs of the forgery levels, it can be seen that the hypotheses validated for the inner security is valid also for the outer security. The results show clearly that the male volunteers produced handwriting samples, which are harder to copy than the females. On all forgery levels, a gap of security is obvious. As it can be seen, the new formulated blind meta attack appears to be an efficient attack especially on handwriting data of women. In general, except for the word *communication*, the women's handwriting data seems to be more endangered to be copied by a meta attack than the male's handwriting data. Background information about writing habits and user information enables the attacker to create effective forgeries. Regarding the security rates of table 6, following samples should be used for handwriting based user authentication.

Table 7: Recommendations for gender based sub groups

Sub groups	**Random**	**Blind**	**Blind Meta**	**Low Force**	**Brute Force**
Male	Symbol/ Signature	Symbol	Sentence	Signature	Word
Female	Short word	Word	Word/ Sentence	Short number/Word	Short word

3rd hypothesis

The third hypothesis is based on the newly established conditional metadata. Comparing the handwriting data, it was obvious that the German users differ in their handwriting habits to each other, this was especially true in case of numbers. Some volunteers wrote the number *1* as a stroke downwards, as shown in figure 3. Others wrote it with a small stroke upward followed by a strike downwards. The survey of the metadata showed that some of the volunteers that handwriting showed the features of figure, whereas those who haven't spent any time abroad mostly wrote as figure 4.

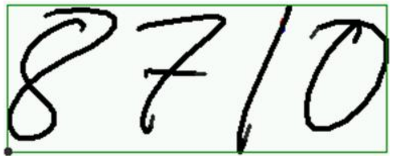

Figure 3: One version of writing number *1*

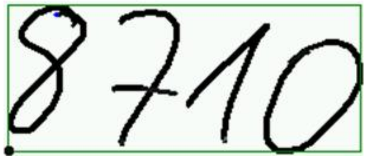

Figure 4: One version of writing number *1*

Because of this conditional metadata background following third hypothesis could be formulated:

Handwriting data of German volunteers having experiences abroad show distinct features that differ from those that lack these experiences. These differences affect especially written numbers.

The evaluations of this hypothesis were applied on the German and English handwriting samples of the German users. To enable inter-linguistic comparability mainly numbers, the individual symbol and signatures were compared. Regarding to the results shown in table 8, the inner security issues of the two sub groups show distinct differences at the given numbers of *8710* und *77993*. The sub group of no experience abroad has better results (EER(8710)=0.20, EER(77993)=0.10) than the opponent group (EER(8719)=0.25 and EER(77993)=0.16). On the other hand signature issues show that peoples having had stays abroad have a more secure handwriting for authentication than those who have not (EER(signature)=0.05 having stays abroad, EER(signature)=0.13 without). On data of given words and sentences in the English language no distinct results can be applied. The hypotheses 3 can therefore be confirmed.

Outer security analysis showed similar characteristics at the numbers. Especially offline information of the low force forgery enabled attackers to copy handwriting aspects of people with experiences of stays abroad sophisticatedly. This sub group, as it is shown in table 8 in general is easier to forge than it's opposite. It is obvious that the differing visual appearance based on differing experiences is an aspect of security for this sub group formed by the metadata of experiences of stays abroad.

Table 8: EERs for stays abroad sub groups

	EER Random		EER Blind		EER Meta		EER Low Force		EER Brute Force	
Semantic	**Abroad**	**Not abroad**	**Abroad**	**Not abroad**	**Abroad**	**Not abroad**	**Abroad**	**Not abroad**	**Abroad**	**Not abroad**
Signature	**0.05**	**0.13**	**0.04**	**0.06**	**0.09**	**0.03**	**0.16**	**0.10**	**0.25**	**0.15**
Symbol	0.05	0.07	0.05	0.06	0.10	0.15	0.20	0.20	0.21	0.23
8710	**0.25**	**0.20**	0.28	0.18	0.30	0.20	0.28	0.20	0.20	0.25
77993	**0.16**	**0.10**	**0.12**	**0.07**	**0.11**	**0.07**	**0.15**	**0.10**	**0.15**	**0.16**
Bird	0.04	0.07	0.06	0.07	0.05	0.03	0.12	0.01	0.08	0.08
Communication	0.06	0.06	0.05	0.05	0.16	0.01	0.10	0.06	0.07	0.08
Where are you going?	0.05	0.03	0.04	0.13	0.06	0.03	0.16	0.15	0.15	0.13

As can bee seen from table 8, not only for numbers, but also for the signature, the lack of security for persons with experiences abroad, can be shown at the evaluation regarding outer security aspects of forgery. On this survey following recommendations for the two groups of experiences abroad and without can be formulated.

Table 9: Recommendations for stays abroad aspects

Sub groups	**Random**	**Blind**	**Blind Meta**	**Low Force**	**Brute Force**
Abroad	Words/Symbol	Signature/ Word	Symbol	Word	Number
Not abroad	Sentence	Word	Signature/ Sentence	Number	Word/Sentence

In order to summarize all recommendations, the following argumentations could be achieved. As can be seen the sentence is the most regarded sample for all user groups. The longer the alphanumeric samples the better is its security issue. Especially at the thread of blind meta forgery sentences give the best security issues. The signature as the classic handwriting sample gives good base for security issues also, but can not be recommended as often. The symbol, a sample that has not been researched before for its security potential in biometrical user authentication also could be recommended. As can be seen great potential for new security issues could be revealed and should be researched in detail in further surveys.

5. CONCLUSIONS AND FUTURE WORK

Based on various metadata, different user groups could be formed and hypotheses were defined, which describe differences between these user groups. The hypotheses could be confirmed to a large portion by the evaluations. It could be shown that the great amount of variability between the EERs of different groups and samples was existent. Therefore distinct recommendations for differing sub user groups could be formulated. Astonishing the fact that the overall used signature could not be recommended as most secure sample for all user groups. Most of the recommendations derived from our experiments consisted of long words, sentences and long numbers. Apart from that, the so far rarely used semantics of symbols seems to hold a special potential for handwriting verification systems. It offers in addition the advantage that it also can be used by illiterates for the authentication with a biometric handwriting system.

For more distinct results, the collected handwriting data should be analyzed completely referring to the collected metadata. The individual and creative samples could be compared to the various given samples. Given samples of the alphabet can be compared to the characters of other given or creative samples. More handwriting data from volunteers of intercultural backgrounds should be researched. Especially the new conditional metadata could be considered and hypotheses about these facts could be formulated. Therefore the outer security aspects could be evaluated in particular using the new blind meta forgery. Then intercultural impact can be achieved and eventually will lead to broad substructures of online handwriting based user authentication.

Based on the results introduced here, a wide field for further research is offered for the future. Presently further group formations and hypotheses creations are in progress. Further the metadata and methodologies introduced here shall be applied to the speech data in future, which have already been taken by the German, Indian and Italian partners. It will be examined by means of metadata, whether speech semantics exist, which are better qualified as other for speaker authentication. Further a multimodal biometric system shall be developed basing on speech and handwriting, which takes into account the different influences by the metadata.

The tests shall further be examined also with the handwriting reference algorithms, which were developed by international research groups in the context of the European project BioSecure. The different nature of the three reference algorithms based on statistic, structural and HMM approaches. A comparison of the results of these different approaches and their multialgorithmic fusion concerning several user groups based on metadata would be a very interesting task in future.

ACKNOWLEDGMENTS

This publication has been produced partly with the assistance of the European Union (project CultureTech, see [7]). The work described in this paper has been also supported in part by the European Commission through the IST Programme under Contract IST-2002-507634 BIOSECURE. The content of this publication is the sole responsibility of the University Magdeburg and their co-authors and can in no way be taken to reflect the views of the European Union.

REFERENCES

1. S. Srihari, C. Huang, H. Srinivasan, *Search engine for handwritten documents*, In: Proceedings of SPIE-IS&T Electronic Imaging, San Jose, California, USA, Vol. 5676, 2005, pp. 66 – 75
2. C. Vielhauer, T. Basu, J. Dittmann, P.K. Dutta, *Finding Metadata in Speech and Handwriting Biometrics,* In: Proceedings of SPIE-IS&T Electronic Imaging, San Jose, California, USA, Vol. 5681, ISBN 3-00-015548-1, 2005, pp. 504 – 515
3. C.I. Tomai, D.M. Kshirsagar, S. Srihari, *Group Discriminatory Power of Handwritten Characters*, In: Proceedings of SPIE-IS&T Electronic Imaging, 2004; pp. 116 – 123
4. T. Scheidat, F. Wolf, C. Vielhauer, *Analyse biometrischer Handschriftverifikation im Kontext von Metadaten*, to appear in Proceedings of GI Sicherheit 2006, Magdeburg, Germany, 2006 (in German)
5. F. Wolf, A. Oermann, C. Vielhauer, P.K. Dutta, T.K. Basu, B. Yegnanarayana, *A Cross-Cultural Evaluation Framework for Behavioral Biometric User Authentication*, In: Proceedings of the 29th Conference of the German Classification Society (GFKL), Special Track on Information Management for User and Data Authentication in IT Security, Magdeburg, 2005
6. P.J. Phillips, P. Grother, R.J. Micheals, D.M. Blackburn, E. Tabassi, M. Bone, *Face Recognition Vendor Test 2002 – Overview and Summary*, 2003
7. The Culture Tech Project, *Cultural Dimensions in digital Multimedia Security Technology, a project funded under the EU-India Economic Cross Cultural Program*, http://amsl-smb.cs.uni-magdeburg.de/culturetech/, requested July 2005
8. C. Vielhauer, R. Steinmetz, A. Mayerhöfer, *Biometric Hash based on Statistical Features of Online Signatures*, In: Proceedings of the IEEE International Conference on Pattern Recognition (ICPR), Quebec City, Canada, Vol. 1, ISBN 0-7695-1696-3, 2002; pp. 123 - 126
9. A.K. Jain, S.C. Dass, K. Nandakumar, *Can soft biometric traits assist user recognition?*, In: Proceedings of SPIE Defense and Security Symposium, Orlando, 2004

10. C. Vielhauer, R. Steinmetz, *Handwriting: Feature Correlation Analysis for Biometric Hashes*, In (Bourlard, H.; Pitas, I.; Lam, K.; Wang, Y., Eds.): EURASIP Journal on Applied Signal Processing, Special Issue on Biometric Signal Processing, Hindawi Publishing Corporation, Sylvania, OH, U.S.A., ISSN: 1110-8657, 2004; pp. 542 - 558
11. C. Vielhauer, *Biometric User Authentication for IT Security: From Fundamentals to Handwriting*, Springer Science+Business Media Inc., New York, USA, ISBN: 0-387-26194-X, 2005
12. F. Zöbisch, C. Vielhauer, *A Test Tool to support Brut-Force Online and Offline Signature Forgery Tests on Mobile Devices*, Proc. of IEEE International Conference on Multimedia and Expo 2003 (ICME), Baltimore, U.S.A., Vol. 3, pp. 225-228, 2003