CURRENT CHALLENGES FOR IT SECURITY WITH FOCUS ON BIOMETRY

BENJAMIN TAMS, MICHAEL TH. RASSIAS AND PREDA MIHĂILESCU

Abstract. In this paper we give a survey of biometrical applications in security context. We start with a brief overview of the different biometric modalities which are most frequently used and compare their security contribution with classical cryptographic primitives. We then consider the case of fingerprints when used as password surrogates. We discuss the main security concerns of biometry in more detail on this practical example and make a point that the *false accept error probability* should be considered as the de facto measure of security.

1. Introduction

Confidential communication is a request with an old tradition, mostly with military applications. Two parties wish to communicate in such a way that no unauthorized (by them) third party may have *a slight chance* to reveal the content of the communication. Some side-requirements in such a setting are

- The request for *secure authentication*.
- The request for provable *signatures*, or, more generally, insurance of the impossibility to repudiate the origin of a message.

A common answer to these requirements was provided by cryptography. A logical art for dealing with this problem is known from early Antiquity; until recent times. It was commonly accepted that for confidentiality, one needed some *secret keys* that were shared only by the authorized parties. The algorithm by which these secret keys were used should also preferably contain some private tricks to make it more reliable. Since the ideas for encryption were based on a common collection of techniques, one could not require completely private algorithms; but it was assumed that by adding some special tricks and complexity, an algorithm would become more resistant to attacks. The general attitude in this respect was completely reversed in modern cryptography, and since decades we prefer to use publically known algorithms, that

Date: Version 1.00 September 3, 2014.

have resisted the scrutiny of a world-wide community of specialists, thus proving their reliability. It is believed that additional private tricks can often lead to providing a false impression of security, which may lead to errors and attacks.

Transposing the alphabet of a spoken language into a sequence of numeric codes is always useful for discussing cryptographic ideas. Suppose thus that the latin alphabet $a, b, \ldots, z$ is encoded in ascending order by the numbers $0, 1, \ldots, 25$. The idea of permuting the letters cyclically by a constant σ was purportedly used by Caesar in the Gallic wars – hence the name of *Caeasar* code. For instance, for $\sigma = 4$, the word

$$ATHENS$$

becomes

$$EYLIRW.$$

For decryption, use $\sigma = 25 - 4 = 21$. The main idea of this approach

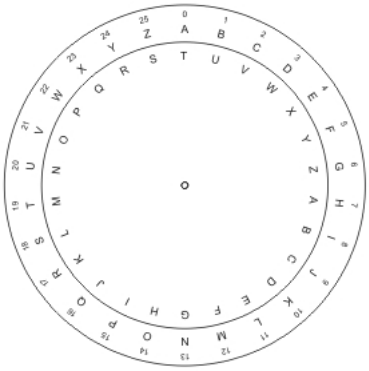


FIGURE 1. Caesar code.

to confidentiality, which is based on the sharing of a secret key – hence the term *secret key algorithms* – is some kind of key triggered permutation. One may permute the very alphabet in which the message is written – the seminal idea used in the so called Caesar Code. More sophisticated variants will first translate the written text by means of a code, and then use chains of key-driven permutations for encryption. This approach is applied even by modern secret key algorithms such as the internationally used *data encryption algorithm* DEA. While *confidentiality* is obtained by protecting the secret key, the authenticity of the peer is only deduced from the fact that he posseses the secret key which should not be obtained by any other person. And there is no possible means to bind messages to the identity of their issuer in secret key mode - by the very fact that at least two peers must share the same

key for communication, it becomes obvious that there is no individual information available for identifying the author of a message.

Therefore, secret key cryptography may still be used successfully even in modern computer times, for protecting the message's content. At the advent of computer networks, the alternative of *public key cryptography* was invented independently by two groups of young american researchers involved in the encipient ARPA net of the 70-es and by an engineer working for the MI5, who was allowed to disclose his discovery in the year 2000. The common idea is to split the key of a peer, say A, for Alice, in two parts, a private part $S(A)$ and a public part $P(A) \subsetneq S(A)$, which is available to the world. Encryption does work both ways like in the secret case. Only, for writing to A, peers B will use the public key, creating messages that can only be decrypted using the private key $S(A)$. In addition, Alice will now be able to authenticate herself, by encrypting *any public message* – for instance "hello world" – with her private key. Since nobody else should be able to find the private key of Alice, upon description with the public key, Bob or anyone else, can be convinced of the fact that it was indeed Alice that generated the encryption – in this situation, the encryption with the private key stands for something like a hand-written signature; it is therefore also called *digital signature*. The same procedure is used to obtain non repudiation of the origin of a message. These facilities are essential for private secure electronic communication.

Therefore public key cryptography has found its exponential spread at the time of the opening of the world wide web to the large public, in the mid-nineties. While public key cryptography is found in more and more applications, several new important problems arise

1. The *reliability* of public keys obtained in the public domain.
2. The multitude of secret key protections required.
3. The very reliability of hardware used in transactions that require personal authentication.
4. User friendliness.

The purpose of this paper is to discuss these challenges of modern IT security. We shall focus hereby on a new facility which receives increasing attention and use in this context, namely the use of biometric traits for identifying humans. After explaining the context in which new challenges to information security arise and discussing the possibilities and limitations of cryptography, we give a brief introduction to the classical aspects of biometry, related to image identification. After that we approach the core subject of this survey, which is the application of biometry to secure applications, giving an overview of

attempts that have been done in this direction, their limitations, and discussing some new algorithms that circumvent problems and vulnerabilities found with some state of the art algorithms.

2. Trends and challenges in information security

In the last three decades, *cryptology* has become a major field of research, together with its Janus - faced duality: *cryptography*, for the design of algorithms and protection principles and *cryptanalysis* for investigation of possible attacks against these algorithms. The primitive algorithms are divided into:

A. Secret Key Algorithms
B. Public Key Algorithms
C. One way functions, hashes and
D. Key management.

We have already discussed briefly the first two. One way functions or hashes have the paradoxical property of being highly non-injective maps, since they map the realm of all possible messages to fixed length blocks, of, say, 192 bits. Such a hash would be a map $\chi : \mathbb{N} \to \mathbb{Z}/(192 \cdot \mathbb{Z})$. However, the size of the image set is large enough to ensure that it is not computationally feasible to find even one collision, i.e. $x \neq y$ with $\chi(x) = \chi(y)$. Little to say about a match, which would require to find, for a given hash of an unknown value, say $h = \chi(x)$ a value $y \in \mathbb{N}$ with $\chi(y) = h$. The collision problem is easier, since it only requires two random hashes to match; in the second case one hash value is already fixed. One way functions must fulfill certain properties related to the conditions discussed. If they do, they are used both for saving passwords in a protected way, without the use of encryption: just substitute a password by its hash value, so that the stored data will reveal no information about the initial password. Hash words are also used in connection with *digital signatures*: Messages to bind to a digital signature are sometimes very large, so one prefers to replace them by their unique hash value and place a digital signature on this hash value.

Key management is less of a cryptographic primitive and more of a set of requirements for the privacy and reliability of keys and passwords used in secure communication. Key management draws on standards of key-authentication, as well as hardware token such as chip cards or other devices, carrying sensitive keys, etc. It is the task of key management to provide not only for secure key storage – either on encrypted memory or chip cards or similar devices – but also for *trust diffusion*. By this we mean that two peers, say Alice and Bob, who

start communication by exchanging public keys, should be provided with means to trust that the received public key does indeed belong to either Alice or Bob. Avoiding attacks by *masquerading* false keys is thus an important task of key management. The provisions for this task are a mixture of cryptography and protocol administration.

It is probably the most important achievement of modern cryptography, that the problems of secure information exchange have been reduced to primitives, endowed with well defined properties, and security is asserted on the base of such properties which can be verified by the cryptologist in the whole world. Hence, the possibility of attacks to a cryptographically secured environment can be also grouped in types of attacks based on well defined *attack scenarios*. It is the presence of these attack scenarios which helps establish the trust into cryptographic solutions, which end up being standardized and used world-wide. A typical, very important *standard* in this context is the *TLS/SSL standard*, which is the cryptographic standard of the world wide web and provides secure communication facilities based on variable tool-kit primitives.

One may conclude that the first decades of public key cryptography provided a reliable system of well scrutinized primitives for addressing each of the problems A. - D. The algorithms for public key encryption, hashes and secret key algorithms as well as the protocols for key management of the last decades are resistant to direct attacks, beyond reasonable doubt. [1]

At the present day, cryptology offers protocols and primitives that are

C1. **Reliable**: They are well researched and secure within any reasonable doubt.

C2. Providing **scalable security** in the sense that it is possible in any of the primitives, to adapt to increased performance of computers, by modifying the length of keys in such a way that the expected time necessary to perform well defined attacks on a given primitive stays unchanged.

C3. **deterministic** in the sense that on the same input they will always produce the same output. The notion of security if based on the provision that an attack on a primitive should require

[1] We should not mistake *beyond reasonable doubt* with *provable certainty*. There is no mathematical proof for the lack of efficient attacks to the state of the art primitives, and even if such one would be provided, it would always be connected to a fixed context of application. But new attacks can be invented, which were not thought of. Confidence relays on the intensive long time research in the public academic domain, spent on the related cryptologic questions.

computation time which stretches beyond hundreds of years, under the most favorable circumstances and using the best algorithms to date. *Even the lowest accepted level of security is beyond doubt*, and the primitives are rejected as soon as theoretical advances show any vulnerability allowing for attacks which can be performed in less than decades or even centuries.

2.1. **Recent evolution.** After these achievements were completed in the 90-es, the challenges of security moved to more volatile topics. The most important ones being:

H1. The definition of trust: in an open environment, who should security protect against whom? Can one trust the user more or the vendor providing some token or hardware, that requires secure identification, which may be stolen?

H2. Viruses and denial of service attacks: both are attacks against an operation system that can either spread over the whole internet or focus on certain target intranets, leading to a blockage of their functionality by overload.

H3. Hacker intrusions of intranets. These are often performed with the purpose of commercial espionage and use any kind of vulnerabilities of operating system, security implementations of even individual authentic users of the intranet.

Developing countermeasures to these very real and corrosive kinds of attacks is an endeavour that requires all the aparatus of cryptology but reaches well beyond: it is the modern task of security engineering.

One may thus observe that cryptology has offered its best and became now part of the more complex task of IT security engineering. Paradoxically, the development and spread of secure applications lead at the opposite end of complexity to new challenges. Since applications are mostly independent and coming from various vendors, the typical user of a large intranet becomes soon confronted with the requirement to secure his identification with respect to a multitude of software, each requiring *safe passwords* from him. This challenges human memory and it mostly happens that users choose to bypass security prescriptions for passwords, by either writing them down or using multiple passwords. This leads to user driven vulnerabilities.

2.2. **The advent of biometry.** In this context, biometry entered the scene by raising an expectation which is best reflected in the paradigm *you are what you are* as opposed to *you are what you know or what you have*. Indeed, in a cryptographical frame, the user is authenticated either by knowledge of some secret, such as the password of some

key, or by means of a token which carries this secret information for him. Assuming she has control on this access modalities, cryptology guarantees secure use. However, the control is relativized by the reasons presented above. Therefore biometry suggests to identify a person physically, by some unique traits that distinguish him uniquely. This can be fingerprints or iris, face or writing mechanics, vein geometry or voice – a multitude of physical and behaviorol traits have been proposed and investigated in order to uniquely identify a person. The wish becomes one to remove the responsibility for identification information from the user and deferr it to technology. The user presents his physical apperance and trusts the system that it may well identify him and not allow intrusions or any other kind of abuse of information related to him. The approach was modivated by the success achieved in image processing during the previous decades, which made the identification by means such as fingerpring, iris or face recognition quite reliable.

However, the advent of biometry and its increasing **actual use** in

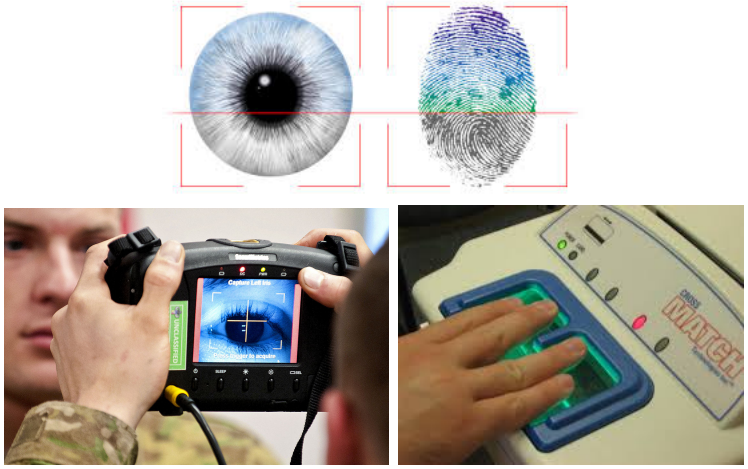


FIGURE 2. Iris and fingerprint recognition.

security contexts raises a series of important questions.

B1. Unlike cryptographic authentication, the biometric one is not deterministic. It is based on comparing real time acquired data - such as images of fingers, iris or a face, against *templates* stored in a database. It is certainly likely that the new data best mactches the one of the template of the individual stored in the data base; this is however only true within a certain stochastic measure of doubt. The actual deterministic certainty is replaced by an error distribution of false acceptances and false

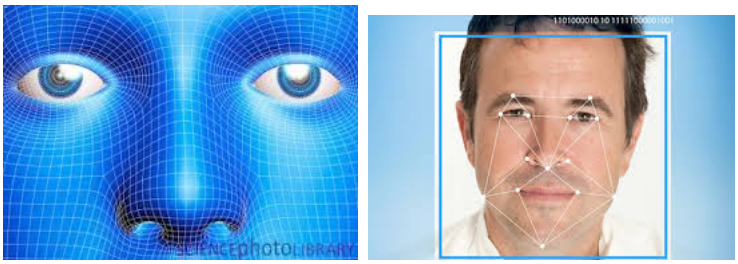


FIGURE 3. Face recognition.

rejects. By deciding acceptance scores, the system can adjust false accepts against false rejects – it will not be able though, to remove any probability of error. This is the **stochastic nature** of biometric identification.

B2. Unlike passwords, which can be replaced when compromessed, biometric traits cannot be changed. As a consequence, the world wide system using one kind of biometry cannot be reliably factored into more secure areas, by use of advanced and expensive technology: a template acquired in a weak system can be used for impersonating a user in any other system.

B3. Most important, the presence of a non vanishing probability of false acceptance becomes the de facto measure of **active information entropy** present in some type of biometric recognition. We describe in this paper some attacks we performed, which confirm practically what one can well imagine by common sense: in presence of a certain probability of false acceptance, one can use data bases of templates for successfully impersonating a stranger.

Although these concerns raise serious caveats for the use of biometry, its user friendliness leads to a continuous propagation of the idea of using it in secure applications. Certainly, the concerns are known, but the vague idea of *application* of lower or medium security concerns was brought as an argument. This breaks the fundamental principle C3. of security: suddenly one seems willing to accept secure contexts, in which attacks can be performed in very short time, yet expecting that the outcome is not sufficiently important for motivating such attacks. It is a defendable point of view, when the security context is a small intranet in which users are satisfied with a formal protection; or when biometry protects access to some protected areas or institutions. However, the consequences in view of B2., namely the uniqueness and irreplaceability of biometric traits are poorly thought through.

As an alternative, a separate branch of activity has been dedicated to a mixture of cryptography and biometry, in which cryptography is supposed to well protect the templates of biometry. However, this quite theoretical area of research operates with the questionable notion of biometric traits being *public data*. This assumption does take into account B2. and the possibility of compromising biometric traits – it is though questionable, what the overall amount of security based on public data may be. Most problematic is the fact that despite intense work, the possibility to uncouple biometric from cryptographic security in these settings, and thus breaking the weakest part in the chain has received no convincing answer yet.

Another approach, which we shall discuss in more detail in this paper, considers biometry as some kind of passwords. They allow access to resources, and, like passwords, should be stored under some one way transformation. This works without problem in the deterministic context: the user presents a password, and its hashvalue is stored. The hash value for one specific password will always be the same, but an attacker cannot recover the password from knowing the hash. In biometry though, any transformation of the template that can allow both to hide the data from intruders and to perform identification, will lead to a notable loss of accuracy in the identification process, as compared to identification by means of "cleartext templates". Therefore, even in the context in which clear template matching provides quite low entropy and thus protection levels, the requirement for password protection leads to an additional loss of entropy, and thus even lower security.

These questions are actively discussed in the literature of the last decade. However, the community of biometry security is a mixed one, ranging from engineers with good expertise in image processing and practical implementations of biometrical matching systems, to specialists of information theory and cryptology who bring new ideas from their domains, while treating biometry as a black box yielding an amount of entropy. The responsibility that the entropy be measurable and sufficient is deferred to applicants – which often are not trained for establishing such complex measures. In fact no mathematical or statistical stringent definition of entropy can be accurately applied in the context. It is one of the points which shall make in this paper, that the de facto entropy of some biometric template is simply given by the equal error rata of the system, i.e. the ballanced probability of false accepts and false rejects. It is a realistic, albeit quite low quantity. A further concern of the encipient biometric security research should be

the one of giving some accurate definitions of attacks. Like in the case of cryptographic security, these attacks should specify clearly:

A1. What ressources one assumes that the intruder may dispose of.
A2. What advantage the intruder wishes to gain.

After providing a brief overview of the currently most frequently used types of biometric identification, we shall focus on the oldest and most spread fingerprint recognition. We shall discuss in this context more in depth the various concerns listed above and provide some partial answers.

3. Overview of Biometry

In our context, biometry is the scientific domain which is concerned with measurements and images of (parts of the) human body, that are to high extent reproducible and may also practically be used for the identification of individuals. In order to be useful in applications, biometry should enjoy some fundamental properties, like

BM1. Universality, meaning that all potential users should possess this biometric trait.
BM2. Uniqueness, in the sense that the biometric trait is different from person to person, and thus helps distinguish individuals and authenticate them correctly.
BM3. Permanence, meaning that the trait will not change in time, and thus, an individual can be identified even on base of templates gathered long periods of time before.
BM4. Some practical properties, such as performance, acceptability and lack of circumvention. The processing time for identification should be low for reach 'acceptable' recognition rates. The acceptability addresses a subjective, social issue: it should be accepted by the bulk of society that presenting one's biometric traits is acceptable. For instance, in some culture, showing the face of a woman and taking pictures of it, might appear as unacceptable, and even presenting one's eye into a camera may require some preparation. Biometric traits may often be imitated by fakes, so it is a requirement mainly for the authentication system, that it be capable of distinguishing between artificial fakes and living biometric sources.

3.1. **Fingerprints.** It was established already before the turn of the last century, that the fingerprints of humans contain sufficient information for distinguishing between any two individuals. Since, in addition, human leave everywhere there fingerprint, due to the sweat and

skin fat, the fingerprint became an important identification method in forensics: techniques for gathering *latent fingerprints* from crime sites developed, together with the science of *dactyloscopy*, which is the craft of fingerprint recognition, in the practice. The fingerprint can be seen as an overall picture of a flaw of ridge lines, induced, in detail with natural endings and bifurcations of the lines. These are called *minutiae* – while cores and delta, visible points of maximal curvature, respectively of divergence of the ridge flaw, are in general easily identified and used for orientation of fingerprint images and templates. George Dalton classified the types of ridge flaws in five main types: left and right loops, whirls and arches, which may be plain or tented. Experience shows that both fingerprints help distinguish even one eyed twins, and the combination of types for the ten fingers is also highly individual. Therefore a first step in matching fingerprints out of large data bases will always begin with a matching of the 10-tuple of types of the ten fingers. This will lead to a small selection within which a detailed identification based on matching of minutiae can be performed by the specialized dactyloscopist. It is agreed that a reliable matching of between twelve and eighteen minutiae is an acceptable base in court, for acknowledging the identity of a person. The precise number of identical minutiae may vary slightly from country to country, and one may even encounter some other classifications of ridge flaws than the one of Dalton - but the main features are the same.

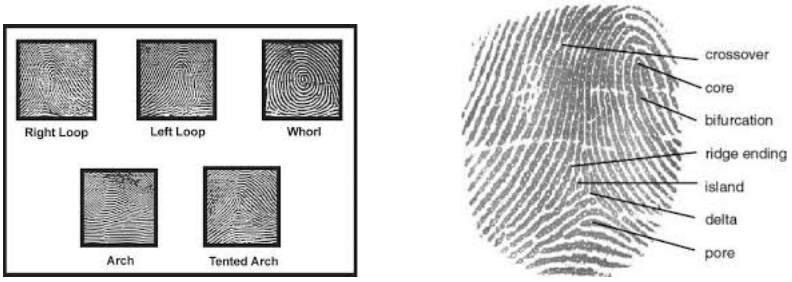


FIGURE 4. Distinguishing fingerprints.

With the advent of computers, the machine - identification of fingerprints became a task of study in image processing; dedicated methods were developed and towards the turn of the century sufficiently reliable plaintext matching system had been developed. For very good quality pictures, an error rate of around 0.1% is frequent, while for pictures of poor quality, even an accuracy of $0.5 - 1.0\%$ is acceptable in practice. Encouraged by the improving quality of matching, the idea of applying

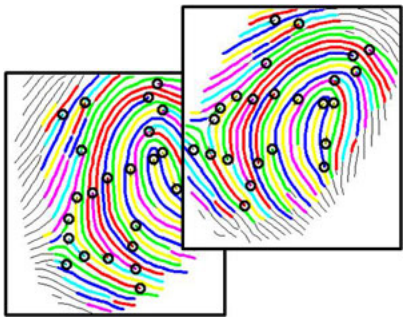


FIGURE 5. Fingerprint matching.

the password paradigm to biometry was brought in the field, first by A. Juels and M. Wattenberg [1] and then again by A. Juels and M. Sudan [2,3]. In this paradigm, biometric templates - fingerprint or others - should be stored in the hashed way, and identification should happen on base of hash values. While in deterministic mode, this approach is very natural, the stochastic nature of biometric matching poses series of problems, and the invention of fuzzy vaults in [2, 3] was the most successful approach for satisfying this requirement. This comes however together with a loss of matching accuracy that may pose serious problems and leads to a difficult decision problem, pondering security against accuracy of matching. As mentioned above, the last is, in the end, also a matter of security – since one must estimate the entropy by the de facto error rate of the system, so when the error rate increases, the security drops too. We shall discuss these issues in more detail in the following chapters.

3.2. **Iris.** While fingerprint recognition has an old, forensic born history, the identification based on the human iris is a one-man show. It was the mathematician John Daugman, presently teaching at Cambridge, who recognized the identification potential in the human iris and developed after a lot of work the algorithms and patents for turning this insight into a practical biometrical identification procedure. The human iris has the advantage of a perfect crown-circular geometry, making its localization in images an easy task. The base for recognition are a system of log-like lines which are different in thickness and frequency, from person to person. Daugman had the bright idea of performing plain Fourier transforms on the iris picture, after having processed it and enhanced image qualities, while unfolding the circle along a line. The result of the analysis is a code of 256 which was standardized and patented by Daugman, as the *iris code*. Claims are

that between $20\% - 30\%$ of identical bits in this code helps ascertain the identity of a person with error rates with in the one per million. Iris has been implemented at various airports, due to its claimed accuracy. Since biometry is not a deterministic science, as soon as iris recognition went public and entered scrutiny of various university research groups, new questions were raised, the claimed recognition rate slightly dropped and even the question was raised, if the iris imprint is permanent in time and if it did not chance after diseases and other organic disturbances. After all, the permanence of the fingerprint had been empirically watched in forensics over more than 100 years, while iris identification is only two decades old. Despite of these discussions, iris recognition is certainly among the leading biometric identification resources, and it has possibly the most impressive accuracy among all. On the other hand, fingerprint recognition can easily improve its performance by using multiple finger recognition.

3.3. **Palm.** Palm recognition is a good alternative to fingers, which gained much popularity in the last decade. The identification artifacts are similar to those for the fingers, but the advantages stem from the fact that palms are easy orientable, better protected from scars and optical disturbances which are a source of poor image quality for fingers, and, finally, have a high amount of information.

FIGURE 6. Palm recognition.

3.4. **Face.** Face recognition is an application as old as computers. There are multiple approaches to face recognition, from flat, two dimensional images, to three dimensional simulations gained by the use of multiple cameras and angles of image acquisition. However, the challenges are very high, since face is the biometrics with the highest dynamics - it may vary due to momentary expression, but also to usual changes, such as make up, eyelid enhancements for women, or beard growing for men. As a consequence, an identification error rate below

5% is quite rare for face recognition. The practical applications are less in the authentication context, and more for the identification of faces in moved contexts and real life scenes.

3.5. **Hand veins.** Hand veins are the youngest type of biometric identification method. It has been claimed that comparing hand vein geometry can lead to identification rates much superior to the one of fingerprints, and reaching in the area of accuracy known from the iris recognition. Picture of hand veins can be taken by means of infrared cameras, which became affordable in the last years, due to technological development. Since hand veins are not exposed, the pictures are very stable and uninfluenced by wounds or physical condition of the scanners. These facts speak in favor of deployment of hand veins as biometrical identification method. Unfortunately, producers of vein-scanners have started a new trend by producing also their proprietary, system embedded, matching algorithms. As a consequence, the academic research with hand veins is at most incipient, and encountering practical problems, being reduced to develop own hard- and software.

FIGURE 7. Hand veins identification.

3.6. **Various other biometries.** The above are the most important and widely spread biometric traits used for identification. However, numerous other typical and distinctive traits have been researched, for the purpose of biometric identification. Voice has an important role in applications of telephony and its potential has thus been thoroughly studied. Termoscans of hands or other parts of the body can also be used for identification, as well as can the mechanics of human gait help recognize individuals with a certain accuracy. Inspired by the hand signature, engineers have built special purpose pens which integrate the *writing mechanics* of individuals, while, for instance, writing down their signature. It is then the mechanical plot and not the actual signature which is used for identification. These and other biometric investigation

either have specific ranges of application where they can be of use, or are a matter of pure research. Their identification rate is in general quite poor, in the range of face recognition.

3.7. **Present applications.** In the last two decades, the applications of biometry reached most diverse areas of social life. In several countries, the drivers' license or id card carries a fingerprint for identification. Meanwhile ATM machines using biometry for identification, based either on iris, palm or fingerprint, are used in several, mostly Asian countries. Fingerprints are used as replacement for signatures especially in third world countries with a high rate of illiteracy. The same kind of biometric traits may also be used for access control in hotels, museums, clubs or lounges, as car openers or weapon activators, etc. In the area of surveillance techniques, face and gait recognition naturally play an important role. While the introduction of biometry in international passports is being pushed ahead world wide, it becomes more and more important to achieve some reliable security standard in the domain of biometric applications in security contexts.

4. "Hashes" for Biometry

Cryptographic password hashes are common solutions for storing passwords in a protected form, while enabling verification of genuine users. However, as discussed above, merely relying on a person's ability to reproduce a password in order to verify her authenticity leads to certain problems – e.g., key management. For this reason, biometry came to be considered as an alternative, possibly in combination to passwords. While the requirements for *biometric template protection solutions* are similar to those for user password protection, they are more difficult to achieve: Passwords are deterministic whereas biometric templates, at the contrary, are typically subjected to noise, i.e., multiple matching samples are expected to be different while they also have some reasonable similarity. These differences can be usefully conceptualized as errors. In this way, biometric template protection schemes have been proposed, that combine techniques known from traditional cryptography with techniques from the discipline of *error-correcting codes* to allow error-tolerant verification.

4.1. **The Fuzzy Commitment Scheme.** One of the conceptually simplest approach for generating protected data from biometric templates that allows error-tolerant verification was proposed by Juels and Wattenberg in 1999 as the *fuzzy commitment scheme* [1].

Let $\mathbf{F}$ be a finite field and assume that we are given the decoder of an *error-correcting code* $\mathbf{C} \subset \mathbf{F}^n$,[2] that is a function

$$\mathsf{dec} : \mathbf{F}^n \to \mathbf{C} \cup \{\textsc{Failure}\},$$

for which there exists an integer $\epsilon \geq 0$ such that $\mathsf{dec}(v) = c$ for all $c \in \mathbf{C}$ and $v \in \mathbf{F}^n$, if the *Hamming distance* fulfills $|c - v| \leq \epsilon$.[3]

Enrollment. On enrollment, given a biometric template encoded as an n-length bit feature vector $x \in \mathbf{F}^n$, its cryptographic hash value $h(x)$ is computed. Then, a codeword $c \in \mathbf{C}$ is chosen at random and the offset $c + x$ is computed. Finally, the hash value together with the offset is stored as the *private template* $(h(x), c + x)$.

Verification. On verification, given a query template $x' \in \mathbf{F}^n$ of the (alleged) same user, an attempt for recovering the protected vector x is performed by computing $(c + x) - x'$. If x' differs in no more than ϵ positions from x, i.e., if $|x - x'| \leq \epsilon$, then $\mathsf{dec}((c + x) - x') = c$ due to the error-correcting property of the decoder. If in this way the correct codeword c can be recovered, then the correct feature vector can be computed according to $x = (c + x) - c$. The correctness of the result can be verified by using its hash value $h(x)$. Otherwise, if $|x - x'| \geq \epsilon$, any verification attempt results, with high probability, in FAILURE or, otherwise, the decoder may output another candidate for the correct feature vector; in both cases, the verification attempt is rejected.

Security. If we assume that the feature vectors x are distributed uniformly and independently among all elements from $\mathbf{F}^n$, then the complexity of the operation of recovering the correct feature set $\mathbf{A}$ from a fuzzy commitment $x + c$ is provably of $O(|\mathbf{F}|^k)$ – or the complexity of breaking the hash $h(x)$ [1]. However, it is not realistic to assume in biometric disciplines that the templates are distributed uniformly within the feature space. We will later emphasize that optimistically estimating security using i.i.d. assumptions easily leads to a severe overestimation of effective security (Section 4.4).

Designing Problems. In order for the fuzzy commitment scheme to be applicable for protecting biometric templates of a certain biometric modality, the following conditions have to be fulfilled:

(1) It must be possible to encode biometric templates as fixed-length feature vectors from $\mathbf{F}^n$.

[2]For more details on error-correcting codes we refer the reader to [4] or any other good textbook on the subject.

[3]Here $|\cdot|$ denotes the *Hamming weight* of a vector in $\mathbf{F}^n$, i.e., the number of positions at which the vector has non-zero entries.

(2) The similarity between biometric templates must be correlated with the Hamming distance of their corresponding feature vectors.
(3) An error-correcting code $\mathbf{C} \subset \mathbf{F}^n$ of sufficient size for which there is a known efficient decoder dec must exist.

Encoding biometric templates as fixed-length feature vectors is usually not a big problem for it is possible to adopt the binary representation of the biometric templates thereby working in the field with two elements. However, ensuring that this binary representation allows comparison via the Hamming distance represents one of the main challenges when designing fuzzy commitment-based biometric template protection. Furthermore, a generic consideration of the concept of error-correcting is not sufficient for implementing a practical fuzzy commitment scheme. First, the code must have a sufficient size in order to allow a certain protection against reversibility attacks; second, it is not known whether there exist codes with efficient decoders for arbitrary block length n.

Even though the generic concept of the fuzzy commitment scheme is very simple, yet clever, the design of a certain fuzzy commitment-based biometric template protection may be challenging. In fact, we have to focus on the specific biometric modalities for which biometry hashes is to be implemented. Thereby, we set our focus to fingerprints even though similar but individual problems exists for other modalities.

4.2. **Fingerprints and its Minutiae.** A *fingerprint* is given by the traces that the ridges of a finger leave on a surface. In these modern days, digital scanners can be used (including specific fingerprint sensors) to obtain a digitized image, i.e., the *fingerprint image* of these traces (see Fig. 8(a) for an example). Typically, features are extracted from fingerprint images on which base two fingerprints can be compared. A standardized type of fingerprint features are *minutiae*, i.e., the positions at which a fingerprint ridge ends abruptly or where it bifurcates, i.e., a *minutiae ending* or *minutiae bifurcation*, respectively. Furthermore, these minutiae positions are typically attached with a *minutia angle* (see Fig. 8(b) for a visualization of an example). Given two minutiae feature sets, i.e., two *minutiae templates*, comparison may be performed through the adoption of two-dimensional point registering methods accounting for the minutia angles. For further details as well as for a comprehensive overview on fingerprints we refer the reader to [5].

(a) (b)

FIGURE 8. A fingerprint (a) and its minutiae (b)

Fuzzy Commitment Scheme for Fingerprint Minutiae. In order to apply the fuzzy commitment scheme for protecting a fingerprint's minutiae template it is necessary to find an encoding of a minutiae set to a space of fixed-length feature vectors in which similarity between minutiae sets is reflected by the Hamming distance. The probably simplest approach to extract a binary n-length feature vector from a minutiae template may work as follows:

(1) The fingerprint image is partitioned into n disjoint regions and each region is attached with a unique index varying between 0 and $n-1$;
(2) For the feature vector $x = (x_0, ..., x_{n-1})$ we set $x_i = 1$ if a minutia is contained in the $i-$th grid region and, otherwise, if no minutia is contained in the $i-$th region, then $x_i = 0$.

The above approach is, however, not very eligible for fingerprint minutiae since usable binary error-correcting codes typically must have a block length n that matches a certain form, e.g., $2^m - 1$ for BCH codes. This explicitly and implicitly leads to limitations when designing fuzzy commitment scheme-based template protection for fingerprint minutiae. There exists an implementation of the fuzzy commitment scheme

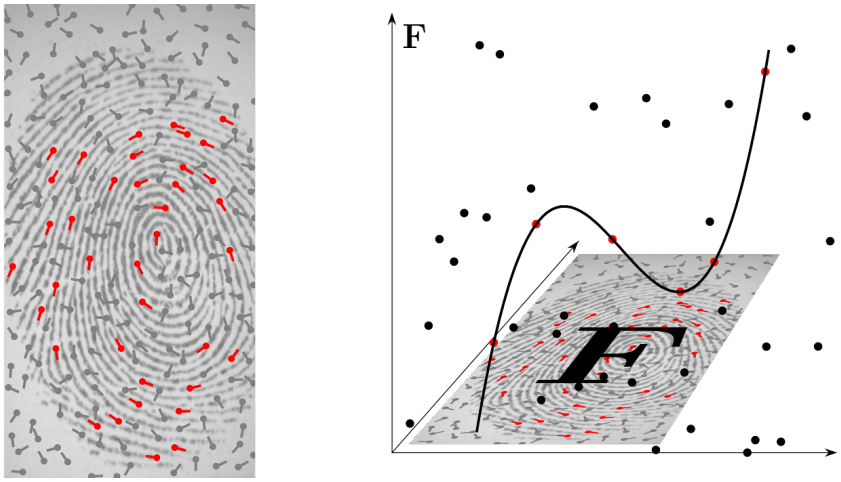
to fingerprint minutiae [6]; nevertheless, the scheme is better tailored for use in other biometric modalities, such as human irises [7].

4.3. **The Fuzzy Vault Scheme.** In 2002, Juels and Sudan [2,3] have proposed the *fuzzy vault scheme* solving some of the problems that we may encounter when attempting to implement a fuzzy commitment scheme for fingerprint minutiae. Like in the case of the fuzzy commitment scheme, the fuzzy vault scheme uses techniques from coding theory in order to conceptualize differences between biometric samples and it has been formulated in quite general terms.

In the following, we shall restrict our considerations to the fingerprint modality. Roughly speaking, the vault works as follows in this case: The minutiae of the to-be-hashed fingerprint, called *genuine minutiae*, are hidden within a large number of randomly chosen non-authentic minutiae, called *chaff minutiae* (see Fig. 9(a)). The genuine minutiae are attached with some information by means of *Reed-Solomon error-correcting codes* while the chaff minutiae are attached with random information deemed to be indistinguishable from the information with which genuine minutiae are constituted thereby providing a certain protection against recovery of the genuine minutiae set from the »minutiae cloud«, i.e., *vault minutiae*. On verification, the minutiae of the query fingerprint are used to extract the *unlocking minutiae*, i.e., those vault minutiae with which query minutiae are of reasonable agreement; if the query fingerprint stems from the same finger, then we may have reason to assume that the unlocking minutiae dominantly consists of genuine minutiae in which case we can recover the entire genuine minutiae set exploiting the error-correcting property; otherwise, if the query minutiae stem from another finger, the unlocking minutiae is expected to contain too few genuine minutiae for recovery.

The eligibility of protecting minutiae with the fuzzy vault scheme has been analyzed by Clancy et al. in 2003 [8] which resulted in a series of minutiae-based fuzzy vault implementations [9–12]. In the following, we present the functioning of a minutiae-based fuzzy vault mainly following the description of Nandakumar, Jain and Pankanti [12].

Enrollment. On enrollment, a minutiae template is given that we want to protect. These minutiae, called genuine minutiae, are mapped to an encoding of a fixed finite field $\mathbf{F}$ by some fixed convention such that there is a one-to-one correspondence between minutiae and the



(a) (b)

FIGURE 9. (a) A fingerprint and its genuine minutiae (red) hidden among a large number of chaff minutiae (gray) (b) Visualization of the genuine minutiae being bound to a Reed-Solomon codeword.

finite field element encoding them.[4] We thus obtain the so-called set of genuine features, also called *feature set* $\mathbf{A} \subset \mathbf{F}$, encoding the minutiae to be protected. We assume that the number of minutiae, namely the size of the feature set, is public and denote it by $t = |\mathbf{A}|$. A secret polynomial $f \in \mathbf{F}[X]$ of degree smaller k is chosen uniformly at random and will be later dismissed. Using f, the genuine pairs $\mathbf{G} = \{(x, f(x)) \mid x \in \mathbf{A}\ )\}$ are computed; this produces a binding of the genuine template to the secret polynomial f. After this, a random set of $n - t$ *chaff features* is generated $\mathbf{A}_{\text{chaff}} \subset \mathbf{F}$ which should be indistinguishable from genuine features $\mathbf{A}$. Finally, chaff pairs $\mathbf{C}$ are generated; they are pairs $(x, y) \in \mathbf{F} \times \mathbf{F}$, in which $x \in \mathbf{A}_{\text{chaff}}$ and y is chosen uniformly among all elements in $\mathbf{F}$ with the constraint that $f(x) \neq y$. The union $\mathbf{V} = \mathbf{G} \cup \mathbf{C}$ finally builds the vault of size n, to which one typically attaches a cryptographic hash $h(f)$ of the secret polynomial, in order to allow secure verification. Consequently, the protected record is given by the pair $(\mathbf{V}, h(f))$.

[4]For example, integral encodings of the abscissa coordinate, ordinate coordinate and angle of a minutia can be concatenated thereby obtaining a single integer that can be used to encode an element of the finite field if of sufficient size. It is important to note that from the finite field element encoded by an integer, the minutia coordinates and minutia angles can be recovered.

Verification. Upon verification, a query feature set $\mathbf{B} \subset \mathbf{F}$ encoding a query minutiae set is provided. Based upon this set, vault pairs are extracted from $\mathbf{V}$, such that the abscissae (encoding a genuine/chaff vault minutia) are well approximated. The unlocking pair set $\mathbf{U} \subset \mathbf{F} \times \mathbf{F}$ is thus determined. If we assume that the query minutiae stem from the same finger as the protected minutiae, then we may expect that a significant amount of query minutiae agree with the genuine minutiae protected by the vault record. In such case, $\mathbf{U}$ may consist of a significant amount of genuine pairs $(x, y) \in \mathbf{G}$, which lie on the graph of the secret polynomial $f \in \mathbf{F}[X]$, i.e., $f(x) = y$. In particular, if $\mathbf{U}$ consists of at least $(|\mathbf{U}| + k)/2$ genuine pairs, then the secret polynomial can be efficiently recovered using an algorithm for decoding Reed-Solomon codes [4].

Brute-Force Security. An intruder who has intercepted a vault record $(\mathbf{V}, h(f))$ may attempt to guess k vault pairs from $\mathbf{V}$ and hope that they are genuine. In case they are genuine, their interpolation polynomial will reveal the correct polynomial of which correctness can be verified with $h(f)$.[5] There are $\binom{n}{k}$ possibilities for an attacker to choose vault pairs of which $\binom{t}{k}$ will reveal the correct polynomial. Hence, with probability $\binom{t}{k} \cdot \binom{n}{k}^{-1}$ an attacker can guess the correct polynomial. This yields a notion of *brute-force security.*

It is important to note that the brute-force attack is based on the unrealistic assumption that minutiae are distributed uniformly and independently from each other. We later emphasize that merely relying on brute-force security as a notion for the security of a fuzzy vault will yield a strong overestimation of the effective security (Sect. 4.4).

Pre-Alignment. A very delicate problem with which implementations of minutiae-based fuzzy vault schemes have to cope with is the problem of fingerprint alignment during a genuine verification process. A common approach is to store unprotected helper data of the protected fingerprint (e.g., points of high ridge curvature) along with the vault records which can be used on verification to pre-align the query templates coarsely in a preliminary step [10–14]. Then, the query minutiae may be adjusted to the vault minutiae to obtain the final alignment with which the unlocking set is extracted [12–14].

[5]Even if the hash were not available, an intruder has still the opportunity to check whether the candidate polynomial interpolates $t = |\mathbf{A}|$ vault pairs; a wrong candidate polynomial will with overwhelming probability not fulfill this requirement for parameters that we expect to encounter in practice, thereby yielding a reliable criteria to an attacker to identify the correct secret polynomial.

From a security perspective, the use of public auxiliary alignment data is problematic because it *does* leak information about the protected fingerprints. Li et al. [15] proposed to use features from the fingerprint that do not depend on the fingers rotation and placement; for instance, features derived from minutiae triangle constellations, thereby removing the issue of information leakage from auxiliary alignment data.

Implementations. One of the first automatic implementation of a fingerprint-based fuzzy vault has been presented by Uludag and Jain in 2006 [11]. They bound the number of minutiae that are protected in the vault by $n = 18$ and bind them to a polynomial of degree smaller than $k = 9$, thereby yielding a brute-force security of 2^{-36}. The genuine acceptance rate that the authors achieved was 73% at which no false accepts have been observed.

In 2007, Nandakumar, Jain and Pankanti [12] improved the genuine acceptance rate to 86% (again, for no observed false accepts) in which at most $t = 24$ genuine minutiae bounded to a polynomial of degree smaller than $k = 11$ are protected within $n = 224$ vault minutiae thereby yielding a brute-force security of 2^{-39}. Nandakumar, Nagar and Jain suggested that the security of their vault implementation can be furthermore improved via a user password [13].

In 2010, Nagar, Nandakumar and Jain showed how additional features of the fingerprint can be used to improve brute-force security by protecting the vaults' ordinate values with a fuzzy commitment scheme [14]. In particular, they showed that a genuine acceptance rate of 92% is achievable at a brute-force security of approximately 2^{-40}.

The above three implementations all require a preliminary alignment step which is supported by public auxiliary alignment data stored along with the vault records. Public data which does leak information can also be exploited by an adversary to improve attacks. Therefore, Li et al. designed a fuzzy vault for fingerprints protecting features that do not depend on the fingerprint's alignment. In this implementation, $t = 40$ genuine features bound to a polynomial of degree smaller than $k = 14$ are hidden within $n = 440$ vault features; this yields a brute-force security of 2^{-52}. The authors measured a genuine acceptance rate of 92%, again at no observed false accepts.

4.4. **Fundamental Security Limit.** Above, the security analyses of the respective fuzzy vault implementations are based on the assumptions that fingerprint features are distributed uniformly and independently from each other in the vault. In this section we give simple but

yet irrefutable arguments why brute-force security is not even a close measure of the implementation's effective security.

We start with a simple exemplary observation. Consider the implementation of Nandakumar, Pankanti and Jain [12]. For the parameter configuration in which $n = 224$ vault minutiae hide at most $t = 24$ genuine minutiae being bound to a secret code polynomial of degree smaller than $k = 9$, the genuine acceptance rate evaluates as 91% while the false acceptance rate estimates as $FAR \approx 0.01\%$. Now, an intruder who has intercepted a vault that he aims to break, i.e., recover the genuine minutiae from it, may establish a large database containing real fingerprints. With these fingerprints he may simulate verification attempts successively until he successfully breaks the vault. Given the computational complexity for simulating an impostor's verification attempt IDT the adversary can expect to break the vault after a time of

$$IDT \cdot \frac{\log(0.5)}{\log(1 - FAR)} \tag{1}$$

yielding the notion of *false-accept security*. In [12] it has been reported that a verification lasted

$$IDT \approx 33 \text{ “Lagrange interpolations”}. \tag{2}$$

Consequently, in terms of Lagrange interpolation for $k = 9$ the false accept security is estimated as approximately

$$2^{18} \text{ “Lagrange interpolations”} \tag{3}$$

which, however, strongly contrasts with an estimated brute-force security of 2^{31} Lagrange interpolations as a realistic measure for the implementation's overall security.

Even in case that no false accepts have been observed during performance evaluation, this does not imply that the false acceptance rate is negligible or even zero: The false acceptance rate is not negligible and the above observation emphasizes more than clearly that **brute-force security is not more than a coarse upper bound for the security of current biometric template protection schemes such as fuzzy vault. Each measure that significantly overestimates false-accept security should be seriously questioned**.

The situation is quite serious. Even a barely usable protection scheme for single finger typically only provides quite-a-low brute-force security of order 2^{31}, say, which is very weak from a cryptographic point of view: It is absolutely no problem to reveal the protected minutiae templates from such a vault within a few minutes. The situation is even worse as an attacker can exploit the statistics of fingerprint minutiae features.

An indication of the maximal achievable security bound is given by the false-accept attack. The complexity of such attacks can be estimated to be in the order of 2^{18}: this amount of operation is a matter of only a few seconds for the attacker—even when using personal computers. In view of these observations, it seems questionable whether there exists sufficient information on a finger in order to achieve a reasonable amount of security. Since methods for fingerprint feature extraction and matching are improving in the process of on-going research, one expects of course according improvements of the security. But even if the false acceptance rate can be reduced to the half – a tremendous improvement, indeed – the false-accept security only slightly improves by a single bit. One can thus hardly expect that fingerprint recognition can evolve in such a way that template protection of single fingerprints may become secure in a cryptographically acceptable way. It is important to note that also for other biometric modalities, such as a human's iris, that can provide higher genuine acceptance rates at lower false acceptance rates than fingerprints, have a security that is still rather low from a cryptographic point of view [7].

4.4.1. *Combination with Passwords.* A possible countermeasure may be to combine passwords with a biometric template, e.g., fingerprint minutiae, to improve security. Such an approach has been implemented and tested in which the minutiae-based fuzzy vault implementation [12] was additionally protected via a 64 bit user password [13]: Using a user password, the vault minutiae are shuffled and on verification given the correct user password, the vault minutiae can be transformed back to their original position. One may argue that the incorporation of user passwords may result in key management problems that were meant to be resolved with biometry: Again, the user of a system has to remember passwords which can be forgotten or, if written down, drop security. On the other hand, biometry can be used to improve password security by a certain amount, for example, 18 bits in case fingerprint minutiae are used—even for easily memorable passwords such as 4-digit *person identification numbers* PINs, say. The weak security of 13 bits provided by a 4-digit PIN can consequently be improved to $32 = 13 + 18$ bits using a single fingerprint's minutiae. For such an approach it must be guaranteed that correctly decrypted vault data is indistinguishable from falsely decrypted vault data—which, in fact, was not guaranteed in [13].

4.4.2. *Slow-Down Functions.* Another approach to improving the low security of biometric template protection is to implement slow-down mechanisms. In a password-based scheme, the password hashes may

be hashed multiple, say a million, times. This yields virtual 20 bits of additional security while also increasing the verification time which is, especially in view of genuine verification attempts, a disadvantage.

In biometric template protection schemes, such as fuzzy vault, merely repeating the hashing process of the secret key's data bound to the template would not be a valid solution. An attacker running a false-accept attack may most likely be able to distinguish the correct secret polynomial from false polynomials without computing its hash. The correct polynomial is of known degree k interpolating $t \gg k$ vault pairs; other query templates will most likely not fulfill this requirement. Similar observations apply to the fuzzy commitment scheme and other constructions based on error-correcting codes [1,16]. Note that this observation has not been accounted for an iris-based fuzzy commitment scheme implementation which was proposed by Hao, Anderson and Daugman [7]. There, the possibility of repeated hashing of the secret codeword has in fact been proposed for improving the security. Nevertheless, this measure yields virtually no additional security due to typically negligible *sphere packing densities*[6] of most *error-correcting codes* [4].

The following may be a valid approach to artificially slow-down the verification process in which the possibility of additionally encrypting the protected biometric templates with password is exploited. A *quiz* $\kappa \in \{0, ..., K-1\}$ is chosen at random during the generation of a protected template and is used to encrypt it. The data of the quiz q is then dismissed. Herewith, the verification process can be artificially slowed-down—in particular an impostor verification attempt. Upon a failing impostor verification attempt, since the correct quiz q is unknown, all possible quizzes $q' = 0, ..., K-1$ must be used to temporarily decrypt the protected reference template and against each temporarily decrypted protected reference template a verification is performed. Consequently, the false-accept security increases by $\log_2(K)$ bits while, on average, the genuine decoding complexity is also increased by a factor of $K/2$. Thus, the slow-down factor K must be chosen carefully in order to achieve sufficient security while still keeping genuine verification feasible. Consequently, the relation between system security and genuine verification time cannot be changed by slow-down measures, i.e., the *security factor* remains unaffected and is potentially low.

4.4.3. *Multiple Fingerprint/Multiple Biometric Modalities.* To overcome the problem of low security factors in a fingerprint-based fuzzy vault, we may consider to fuse multiple fingerprints acquired from a user and

[6]sphere packing density: $|\mathbf{F}|^{n-k} \sum_{j=0}^{\epsilon} (|\mathbf{F}|-1)^j \binom{k}{j}$ where $k = \dim(\mathbf{C})$

protect them with the fuzzy vault scheme. On genuine verification, more than one fingerprint may be required for an accept. On the other hand, breaking a fuzzy vault to multiple fingerprints may also be more secure against attacks, in particular, false-accept attacks. An implementation of a multi-finger fuzzy vault has been proposed by Merkle et al. in 2011 [17]. It is important to note that the implementation has not been evaluated in terms of genuine acceptance rate and false acceptance rate and it is still unclear how a multi-finger implementation can perform.

It is also possible to fuse multiple biometric modalities of which the fusion of fingerprints is a special case. In 2012, Nagar et al. [18] analyzed fusion strategies of fingerprints, irises, and face using the fuzzy vault and fuzzy commitment scheme and reported that it is possible to achieve a 75% genuine acceptance rate at a security level of 53 bits. In principle, this is an interesting result. It remains, however, unclear if and in which applications a fusion of fingerprints, iris, and face, which may be related with several inconveniences for users, will be of interest. Especially under the circumstance that the moderately high security of 53 bits is compensated by the very low genuine acceptance rate of merely 75%.

4.5. **Attacks via Record Multiplicity.** Even if we can assume that it is possible to implement a usable biometric template protection scheme, possibly based on multiple fingerprints or more generally on multiple biometric modalities, there are, however, other risks that must be considered. In addition to mere off-line attack in which an adversary aims at revealing the protected templates from intercepted data, there exist another serious scenario in which an adversary who has intercepted two (or more) protected records attempts to decide whether they stem from the same finger, say, i.e., whether they are *related*. The process of distinguishing related from unrelated records is commonly called *cross-matching* and is a privacy risk with which an intruder having intercepted the content of multiple application's database could trace particular users activity. For this reason international standards explicitly require from biometric template protection schemes to be *unlinkable*, i.e., cross-matching must not be possible (ISO/IEC IS 24745 [19]).

4.5.1. *Correlation Attack in a Fuzzy Vault Scheme.* In general, the fuzzy vault scheme is vulnerable to a very serious cross-matching attack [20]. Observe that in a fuzzy vault, the genuine features stem from a biometric sample while the chaff features have been generated

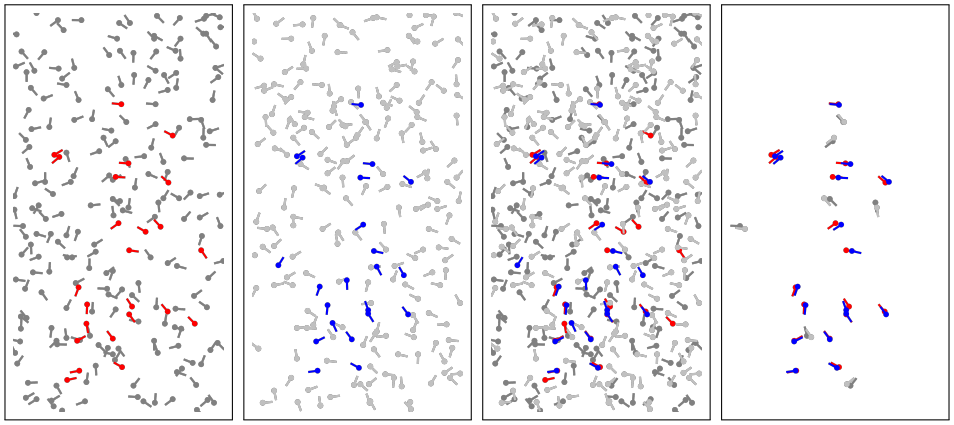


FIGURE 10. Visualization of the correlation attack process in a minutiae-based fuzzy vault: those vault minutiae of two related vaults (a) and (b) are correlated (c) and those vault minutiae that well agree have a quite good chance to be genuine minutiae (d) being colored red and blue for the first and second vault, respectively.

at random. If two fuzzy vault record can be intercepted by an intruder protecting templates that stem from the same instance (e.g., finger) we may observe that the genuine vault features in the first record (e.g., red-colored minutiae in Fig. 10(a)) well agree with the genuine vault features in the second record (blue-colored in Fig. 10(b)), i.e., the correlate well as compared to the chaff features (Fig. 10(c) and Fig. 10(d)). This property can be exploited by an intruder to distinguish related from unrelated vault correspondences. Even worse, an intruder who has intercepted two related vault records may even unlock the vaults using the candidate sets of genuine vault features. In fact, for a minutiae-based fuzzy vault implementation, Kholmatov and Yanikoglu [21] demonstrated that an intruder can break two related vault correspondences with success probability of order 60%, which is much too high for a system to fulfill the unlinkability and irreversibility requirement. The possibility of running the so-called *correlation attack* calls for a valid countermeasure.

4.5.2. *Decodability Attack in a Fuzzy Commitment Scheme.* At a glance the serious vulnerability of the fuzzy vault scheme not to fulfill the unlinkability requirement advocates to base the protection on the fuzzy commitment scheme (see Sect. 4.1). However, the fuzzy commitment

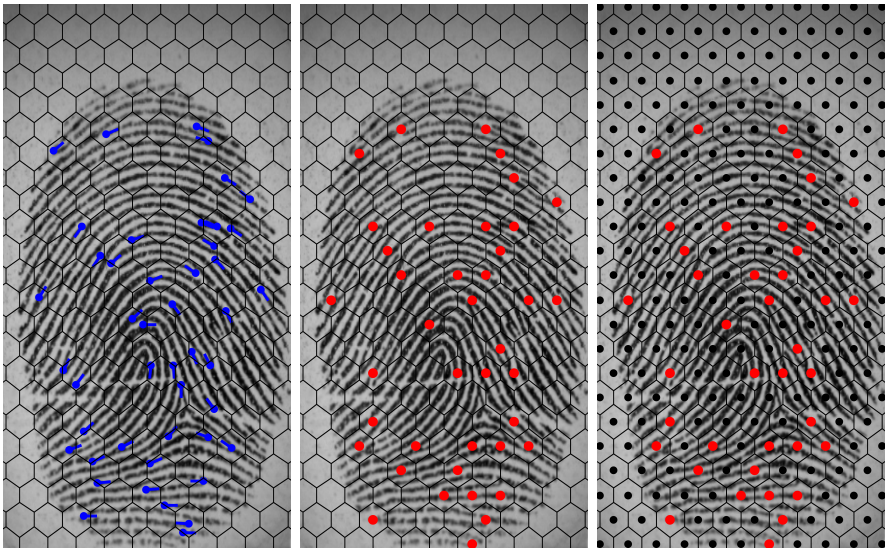
scheme is vulnerable to a linkability attack, too. With the notation of Sect. 4.1, assume that an intruder has intercepted two related records of the fuzzy commitment scheme $c + x$ and $c' + x'$, i.e., where $c, c' \in \mathbf{C}$ are random elements of a linear code $\mathbf{C} \subset \mathbf{F}^n$ and $x, x' \in \mathbf{F}^n$ are feature vectors with Hamming distance within the code's error-correcting capability ϵ, i.e., $|x - x'| \leq \epsilon$. The intruder has the possibility to compute the difference

$$(c + x) - (c' + x') = (c - c') + (x - x') \tag{4}$$

and exploit the observation that $c - c'$ is a codeword, due to the linearity of $\mathbf{C}$, and the bound $|x - x'| \leq \epsilon$. Hence, the difference can be decoded to the codeword $c - c'$ given two related records of the fuzzy commitment scheme. For non-related records, i.e., where $|x - x'| > \epsilon$, the difference may be decodable with probability equal to the sphere packing density of $\mathbf{C}$; this is typically negligible for most linear codes used in implementations of the fuzzy commitment scheme. Thus, just from decodability of the difference, an intruder may distinguish related from non-related records thereby conflicting with the unlinkability requirement. It is known that capturing two related records based on the linear code brings no advantage for breaking the fuzzy commitment scheme. But if an intruder has intercepted two related records based on different linear codes, the irreversibility requirement cannot be guaranteed [22].

In a binary fuzzy commitment scheme, Kelkboom et al. [23] proposed to pass the feature vectors through a record-specific but public permutation process, in order to prevent the decodability attack. Unfortunately, it has been overlooked that by implementing the measure, two related records of the fuzzy commitment scheme being subjected to different public permutation processes can be considered as having been built by means of different linear codes. This makes them susceptible to the reversibility attack mentioned above [24]. It has furthermore been shown in [24] that in a binary fuzzy commitment scheme, the problem cannot be solved by passing the feature vectors through a public transformation process that preserves the Hamming distance between two feature vectors. Fortunately, there may exist such transformations for a non-binary fuzzy commitment scheme. However, most implementations of the fuzzy commitment scheme are used to protect binary biometric feature vectors and thus the problem of designing an effective binary template protection scheme remains a challenge.

4.5.3. *Unlinkable Minutiae-Based Fuzzy Vault.* The correlation attack in a fuzzy vault scheme yields an advantage to an attacker in linking



(a) (b) (c)

FIGURE 11. Visualization of how to make a minutiae-based fuzzy vault resistant against the correlation attack: (a) The genuine minutiae are rounded/quantized as coordinates of a (for example) hexagonal grid (b) and all other unoccupied grid coordinates are used as the chaff (c)

and breaking two related records due to the fact that the chaff is generated at random while the genuine features stem from the same instance thereby essentially being fixed (up to tolerable noise). We may avoid this inconvenience in a fuzzy vault scheme by a simple yet effective variation, which we describe next, in informal terms, for fingerprint minutiae. A grid (e.g., rectangular or hexagonal) is laid over the fingerprint image; each genuine minutiae is rounded to grid coordinates that are then used to build the genuine features, thereby passing the genuine minutia through a quantization scheme (we may also quantize the minutiae angles in a similar manner). All other, unoccupied grid coordinates are used as the chaff. Consequently, there is no correlation that can be exploited in an attack, since the feature sets are equal for any two records.

Some challenges remain with this approach. Upon verification, it is not possible to adjust query minutiae to vault minutiae in order to unlock the vault with a pre-aligned minutiae set. Alternatively, we have to ensure that the genuine minutiae and the query minutiae can be

represented w.r.t. an intrinsic coordinate system that can be robustly extracted from a fingerprint. This introduces new error sources. In fact, the estimation of intrinsic coordinate system and an alignment-free representation of minutiae, i.e., *absolutely pre-aligned minutiae*, is a challenging problem for which no definite solution has been found [5].

Recently, some progress in automatic absolute minutiae pre-alignment has been presented [25] and evaluated for an unlinkable minutiae-based fuzzy vault. The genuine acceptance rate that is currently achievable with such an approach is of order 80% at which no false accept has been observed while in a traditional minutiae-based fuzzy vault the genuine acceptance rate has been measured as 86% on the same database. Consequently, the unlinkability requirement can be fulfilled at a genuine acceptance rate well comparable to a traditional approach which, however, is prone to record multiplicity attacks. It is important to note that both implementations following the traditional approach and by applying a quantization scheme to the minutiae are subject to the fundamental security limit discussed in Sect. 4.4. However, by incorporating a quantization scheme and robust methods for absolute fingerprint pre-alignment we may eventually obtain an unlinkable biometric template protection scheme for multiple fingerprints and/or even multiple biometric modalities. It remains to be seen how implementations for multiple fingerprints will be able to perform regarding verification performance and security.

4.5.4. *A Compact Fuzzy Vault Scheme.* Passing absolutely pre-aligned minutiae through a quantization process has another advantage, beyond merely achieving resistance against the correlation attack. We can apply a modified fuzzy vault construction proposed by Dodis et al. [16] for protecting quantized minutiae sets. This has the advantage of producing significantly more compact record sizes.

As above, the quantized minutiae are encoded as a subset $\mathbf{A}$ of the underlying finite field $\mathbf{F}$. Furthermore, as before, let $f \in \mathbf{F}[X]$ be a secret polynomial of degree smaller than k. Instead of chaff generation thereby yielding a set of vault pairs explicitly, they are encoded by the following polynomial

$$V(X) = f(X) + \prod_{x \in \mathbf{A}} (X - x). \tag{5}$$

If $x \in \mathbf{A}$, then $V(x) = f(x)$ and thus $(x, V(x))$ is a genuine pair lying on the graph of the secret polynomial; otherwise, if $x \notin \mathbf{A}$, then $V(x) \neq f(x)$ and then $(x, V(x))$ is a chaff pair. Consequently, by a single compact polynomial, genuine and chaff pairs are encoded in

a smart manner. Note that $V(X)$ is a monic polynomial of degree $t = |\mathbf{A}|$. It only requires $t \cdot \log_2(|\mathbf{F}|)$ storage bits, while the traditional vault would need $2 \cdot (t + |\mathbf{C}|) \cdot \log_2(|\mathbf{F}|)$ bits for storing the vault pairs explicitly.

On the other hand, the following fact can be shown. Suppose that two related records of the compact fuzzy vault scheme can be intercepted:

$$V(X) = f(X) + \prod_{x \in \mathbf{A}} (X - x) \tag{6}$$

$$W(X) = g(X) + \prod_{x \in \mathbf{B}} (X - x), \tag{7}$$

and these records are protecting the feature sets $\mathbf{A}, \mathbf{B}$ bound to the polynomials $f, g \in \mathbf{F}[X]$ both of degree smaller than k, respectively. Here, without loss of generality, we assume that $|\mathbf{A}| \geq |\mathbf{B}|$ and $|\mathbf{A} \cap \mathbf{B}| \geq (|\mathbf{A}|+k)/2$ is fulfilled. Then the differences $\mathbf{A} \backslash \mathbf{B}$ and $\mathbf{B} \backslash \mathbf{A}$ can be recovered explicitly and efficiently by applying the extended Euclidean algorithm to $V(X)$ and $W(X)$. We refer to [26] for a proof of this fact. This would again conflict with the unlinkability requirement of effective biometric template protection calling for a countermeasure. Fortunately, by passing the feature elements through a record-specific random but public permutation $\mathbf{F} \to \mathbf{F}$ is a promising solution for preventing the extended Euclidean algorithm-based record multiplicity attack [26].

A record-specific random bit permutation process was considered to be incorporated in a fuzzy commitment scheme, in order to prevent the decodability attack [23]. In view of the fact that this measure has been shown to be forgeable [25], it would be highly desirable to prove or disprove the validity of the countermeasure in a reductionist sense, say. Yet, there is currently no attack known that can break two related records of the compact fuzzy vault scheme being subjected to a record-specific permutation process significantly better than breaking one of the records individually.

4.6. **The Future of Biometric "Hashes".** The major issue in providing information security for biometric templates may lay in the design of implementations. In particular, for specific biometric modalities suitable feature extractions have to be developed that can decrease the most limiting factor *false acceptance rate* at a maintained and preferably high genuine acceptance rate. However, even if the false acceptance rate can be reduced to its half for a certain biometric modality which would be a breakthrough, the security only increases by a single

bit. Even though reducing false acceptance rates is certainly worth its trouble, it seems more reasonable to rely on the fusion of multiple biometric systems to achieve an acceptable amount of security. First steps have already been made [18], but they leave space for improvement.

References

[1] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in *Proc. of ACM Conf. on Computer and Communications Security*, 1999, pp. 28–36.

[2] A. Juels and M. Sudan, "A fuzzy vault scheme," in *Proc. Int. Symp. Inf. Theory*, A. Lapidoth and E. Teletar, Eds., 2002, p. 408.

[3] ——, "A fuzzy vault scheme," *Des. Codes Cryptography*, vol. 38, no. 2, pp. 237–257, 2006.

[4] E. R. Berlekamp, *Algebraic coding theory*. Laguna Hills, CA, USA: Aegean Park Press, 1984.

[5] D. Maltoni, D. Maio, A. Jain, and S. Prabhakar, *Handbook of Fingerprint Recognition*, 2nd ed. Springer Publishing Company, Incorporated, 2009.

[6] A. Arakala, J. Jeffers, and K. Horadam, "Fuzzy extractors for minutiae-based fingerprint authentication," in *Proc. Int. Conf. on Biometrics*, ser. LNCS 4642, 2007, pp. 760–769.

[7] F. Hao, R. Anderson, and J. Daugman, "Combining crypto with biometrics effectively," *IEEE Trans. Comput.*, vol. 55, no. 9, pp. 1081–1088, Sep. 2006.

[8] T. C. Clancy, N. Kiyavash, and D. J. Lin, "Secure smartcard-based fingerprint authentication," in *Proc. ACM SIGMM workshop on Biometrics methods and applications*, ser. WBMA '03. New York, NY, USA: ACM, 2003, pp. 45–52.

[9] S. Yang and I. Verbaudwhede, "Automatic secure fingerprint verification system based on fuzzy vault scheme," in *Proc. Int. Conf. on Acoustics, Speech and Signal Processing*, 2005, pp. 609–612.

[10] U. Uludag, S. Pankanti, and A. K. Jain, "Fuzzy vault for fingerprints," in *Proc. Int. Conf. on Audio- and Video-Based Biometric Person Authentication*, 2005, pp. 310–319.

[11] U. Uludag and A. K. Jain, "Securing fingerprint template: fuzzy vault with helper data," in *Proc. Workshop on Privacy Research In Vision*, 2006, pp. 163–169.

[12] K. Nandakumar, A. K. Jain, and S. Pankanti, "Fingerprint-based fuzzy vault: Implementation and performance," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 4, pp. 744–757, 2007.

[13] K. Nandakumar, A. Nagar, and A. Jain, "Hardening fingerprint fuzzy vault using password," in *Proc. Int. Conf. on Biometrics*, ser. LNCS 4642, 2007, pp. 927–937.

[14] A. Nagar, K. Nandakumar, and A. K. Jain, "A hybrid biometric cryptosystem for securing fingerprint minutiae templates," *Pattern Recogn. Lett.*, vol. 31, pp. 733–741, June 2010.

[15] P. Li, X. Yang, K. Cao, X. Tao, R. Wang, and J. Tian, "An alignment-free fingerprint cryptosystem based on fuzzy vault scheme," *J. Netw. Comput. Appl.*, vol. 33, pp. 207–220, May 2010.

[16] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," *SIAM J. Comput.*, vol. 38, no. 1, pp. 97–139, 2008.
[17] J. Merkle, H. Ihmor, U. Korte, M. Niesing, and M. Schwaiger, "Performance of the fuzzy vault for multiple fingerprints (extended version)," *CoRR*, vol. abs/1008.0807v5, 2011.
[18] A. Nagar, K. Nandakumar, and A. K. Jain, "Multibiometric cryptosystems based on feature-level fusion," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 1, pp. 255–268, 2012.
[19] ISO/IEC JTC1 SC2 Security Techniques, "ISO/IEC 24745:2011. Information Technology - Security Techniques - Biometric Information Protection," International Organization for Standardization, 2011.
[20] W. J. Scheirer and T. E. Boult, "Cracking fuzzy vaults and biometric encryption," in *Proc. of Biometrics Symp.*, 2007, pp. 1–6.
[21] A. Kholmatov and B. Yanikoglu, "Realization of correlation attack against the fuzzy vault scheme," in *Proc. SPIE*, vol. 6819, 2008.
[22] K. Simoens, P. Tuyls, and B. Preneel, "Privacy weaknesses in biometric sketches." in *IEEE Symp. on Security and Privacy*. IEEE Computer Society, 2009, pp. 188–203.
[23] E. J. C. Kelkboom, J. Breebaart, T. A. M. Kevenaar, I. Buhan, and R. N. Veldhuis, "Preventing the decodability attack based cross-matching in a fuzzy commitment scheme," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 1, pp. 107–121, 2011.
[24] B. Tams, "Decodability attack against the fuzzy commitment scheme with public feature transforms," *CoRR*, vol. abs/1406.1154, 2014.
[25] ——, "Absolute fingerprint pre-alignment in minutiae-based cryptosystems," in *Proc. of BIOSIG*, 2013, pp. 75–86.
[26] J. Merkle and B. Tams, "Security of the improved fuzzy vault scheme in the presence of record multiplicity (full version)," *CoRR*, vol. abs/1312.5225, 2013, in review.

(B. Tams) Mathematisches Institut der Universität Göttingen, Bunsenstrasse 3-5, D-37073, Göttingen, Germany
E-mail address, B. Tams: btams@math.uni-goettingen.de

(M. Th. Rassias) Department of Mathematics, ETH–Zürich, Rämistrasse 101, 8092, Zürich, Switzerland
E-mail address, M. Th. Rassias: michail.rassias@math.ethz.ch

(P. Mihăilescu) Mathematisches Institut der Universität Göttingen, Bunsenstrasse 3-5, D-37073, Göttingen, Germany
E-mail address, P. Mihăilescu: preda@uni-math.gwdg.de