International Journal of Pervasive Computing and Com

Privacy-preserving biometrics authentication systems using fully homomorphic encryption

Wilson Abel Alberto Torres Nandita Bhattacharjee Bala Srinivasan

Article information:

To cite this document:
Wilson Abel Alberto Torres Nandita Bhattacharjee Bala Srinivasan , (2015),"Privacy-preserving biometrics authentication systems using fully homomorphic encryption", International Journal of Pervasive Computing and Communications, Vol. 11 Iss 2 pp. 151 - 168
Permanent link to this document:
http://dx.doi.org/10.1108/IJPCC-02-2015-0012

Downloaded on: 07 November 2016, At: 22:34 (PT)
References: this document contains references to 51 other documents.
To copy this document: permissions@emeraldinsight.com
The fulltext of this document has been downloaded 640 times since 2015*

Privacy-preserving biometrics authentication systems using fully homomorphic encryption

Wilson Abel Alberto Torres, Nandita Bhattacharjee and Bala Srinivasan
Faculty of IT, Monash University, Melbourne, Australia

Received 28 February 2015
Revised 28 February 2015
Accepted 5 April 2015

Abstract

Purpose – The purpose of this paper is to determine the effectiveness of using fully homomorphic encryption (FHE) to preserve the privacy of biometric data in an authentication system. Biometrics offers higher accuracy for personal recognition than traditional methods because of its properties. Biometric data are permanently linked with an individual and cannot be revoked or cancelled, especially when biometric data are compromised, leading to privacy issues.

Design/methodology/approach – By reviewing current approaches, FHE is considered as a promising solution for the privacy issue because of its ability to perform computations in the encrypted domain. The authors studied the effectiveness of FHE in biometric authentication systems. In doing so, the authors undertake the study by implementing a protocol for biometric authentication system using iris.

Findings – The security analysis of the implementation scheme demonstrates the effectiveness of FHE to protect the privacy of biometric data, as unlimited operations can be performed in the encrypted domain, and the FHE secret key is not shared with any other party during the authentication protocol.

Research limitations/implications – The use of malicious model in the design of the authentication protocol to improve the privacy, packing methods and use of low-level programming language to enhance performance of the system needs to be further investigated.

Originality/value – The main contributions of this paper are the implementation of a privacy-preserving iris biometric authentication protocol adapted to lattice-based FHE and a sound security analysis of authentication and privacy.

Keywords Security, Biometric, Biometric authentication system, Fully homomorphic encryption, Privacy-reserving

Paper type Research paper

1. Introduction

Currently, the digital revolution lets us manipulate plenty of information that has to be available all the time. Consequently, this revolution requires multiple interactions in every field of society, so applying some forms of security to protect these interactions may be required. The forms of this security might be confidentiality, integrity of the information or even privacy protection. Privacy is defined as the right that every individual has to control the collection and use of information about themselves: "the right to be left alone" (Caloyannides, 2003). In the past, government and military applications were identified as requiring the protection of privacy. An example would be classified sensitive information or preventing weapon systems from being activated by unauthorized people. Subsequently, because of the digital revolution, privacy has begun

International Journal of Pervasive Computing and Communications
Vol. 11 No. 2, 2015
pp. 151-168
© Emerald Group Publishing Limited
1742-7371
DOI 10.1108/IJPCC-02-2015-0012

to play an important role in the protection of individuals' private information and also for industry. To address basic security needs, such as privacy and authenticity, encryption and authentication schemes have been developed.

Biometric methods are often presented as a better approach to user recognition. They help to avoid the need to carry tokens or remember passwords to achieve authentication in an access control system because of their properties, such as uniqueness, universality, permanence, collectability and usability (Prabhakar *et al.*, 2003). However, if biometric information is stored in an unprotected manner, it will lead to privacy and security risks. As a result, it is necessary to protect the privacy of biometric information with encryption techniques.

This paper is organized in six parts, including the introduction. Section 2 gives a brief background of biometric authentication systems and some approaches to preserving the privacy of biometric data. After describing the biometric authentication protocol used in Section 3, this research shows how fully homomorphic encryption (FHE) can be implemented in a biometric authentication system in Section 4. Section 5 analyzes this approach in terms of attacks, desirable properties and FHE properties. Section 6 will conclude and suggest directions for further research.

2. Background

Security plays an important role because it guarantees that information is accessed or manipulated exclusively by authorized users. User authentication is a fundamental block in computer security, which verifies the claimed identity of the user. It is used to ensure that only authorized users can access the system. To fortify this verification, Stallings and Brown (2012) mention four methods or means of authentication:

(1) something the individual knows (passwords and PIN);

(2) something the individual possesses (tokens);

(3) something the individual does, also known as dynamic biometrics or behavioral biometrics (voice pattern, handwriting and keystroke); and

(4) something the individual is, also known as static biometrics or physiological characteristics (fingerprint, face, DNA sequences, iris and hand).

Among the biometric traits presented, the most robust is the iris. This is because the iris is formed during the development of the embryo and remains unchanged during the whole lifetime. Also, the iris has enormous pattern variability among different persons, and there is no other genetic correlation. Identical individuals such as twins, or even individual eyes (either right or left), can be easily distinguished. Furthermore, it is very difficult to modify the iris pattern by any surgery (Daugman, 1997). Finally, it is mentioned that iris patterns can be read 300 times faster than fingerprint patterns when running a matching process against a biometric database (Mollin, 2006).

Biometrics, used to recognize individuals through behavioral or physiological characteristics, is a powerful tool against repudiation and is largely immutable because of its properties. The acquired biometrics are sensitive to small changes or variations in their inputs; as a result, biometric data are also referred to as "noisy data". This is for many reasons. For example, humidity variation, temperature and thermal noise in the measurement system or interface that senses or detects the biometrics of the individual (Škorić, 2010). Recognition accuracy is established by using statistical decision theory.

The classical framework of statistical decision theory is based on the binary (yes/no) decision schema offered by biometric systems, as represented in Daugman (2000), which provides four possible results: a false accept rate (FAR), a correct accept rate (CAR), a false reject rate (FRR) and a correct reject rate (CRR). The relative probabilities of these four results can be adjusted in a way that reflects their associated costs and benefits. In practice, it is almost impossible to get both zero FAR and FRR errors because the intra (genuine) and inter (imposter) classes are difficult to completely separate in the measurement space (Daugman, 2000).

Biometric verification systems have two types of errors in their verification steps (Jain *et al.*, 2004):

1. *A false match*: Misrecognizing measurements from two different people as being from the same person, called false acceptance.
2. *A false non-match*: Misrecognizing measurements from one person as being from two different persons, called false rejection.

All biometric authentication systems require two steps: "enrollment" and "authentication". To initiate biometric authentication, first, a user must be enrolled. In this phase, the biometric characteristics are sensed and converted to a set of numbers, also known as a user template, and this user template is stored in a central database. Second, authentication can be used as verification (match one-to-one) or identification (match one-to-many), in which the system extracts the biometric information, then compares it with that stored in the database to authenticate or identify a user based on a given threshold (Stallings and Brown, 2012).

Some of the biggest concerns in biometric systems are security and privacy. Regarding security, where the goal is to protect the biometric data against an unauthorized user, it is shown that some biometric traits are easily exposed (fingerprints, face and voice) and can be stolen for identity theft. Furthermore, the properties of biometrics, permanence and immutability, considered its main strengths, are also a weakness because raw biometric data are permanently linked with the person and cannot be reissued, revoked or cancelled if this information is compromised. Basically, because there are limited numbers of biometric characteristics per individual, for instance two eyes, one face or one voice. Regarding privacy, where biometric data are shared with specific people, individual characteristics represented by biometrics are disclosed, and this can lead to user profiling, discrimination and loss of anonymity (Campisi, 2013).

Commonly, biometric protection templates are categorized as:

- Feature transformation (FT), where the transformation value is only stored in the database, some approaches of this category include biohashing (Goh and Ngo, 2003) and cancellable biometrics (Ratha *et al.*, 2001; Bolle *et al.*, 2002).
- Biometric cryptosystems (BC), mainly implemented to generate a cryptographic key from biometric features or by using the biometric features secured by a cryptographic key, some approaches are fuzzy commitment (Juels and Wattenberg, 1999) and fuzzy vault (Juels and Sudan, 2006). Fuzzy schemes are still a storage-oriented approach and are not secure against indistinguishability and irreversibility attacks (Bringer *et al.*, 2013; Simoens *et al.*, 2009). Besides that, Scheirer and Boult (2007) mentioned that both "FT" and "BC" are vulnerable to record multiplicity, surreptitious key inversion and blended substitution attacks.

Among the approaches described, none of these ensures that privacy is preserved in biometric data.

- Moreover, some approaches that focus on working over an encrypted domain are mentioned in locality-sensitive hashing and bloom filters (Bringer *et al.*, 2011): Anonymous Biometric Access Control (ABAC) (Ye *et al.*, 2009) and secure multiparty techniques (Bringer *et al.*, 2013). It is observed that the common technique used in all of these is homomorphic encryption (HE). HE itself has different versions and, mostly because of its limitations, has been used along with other techniques to achieve good levels of privacy.
- HE, which was mentioned initially in Rivest *et al.* (1978), follows the foundation of public key encryption, such as RSA or ElGamal. The limitation of working with encrypted data using encryption functions before the data are decrypted was exposed at that time; the encryption functions were called "privacy homomorphisms". This method can be explained with an example: imagine that Alice encrypts the input "message" and sends the ciphertext to Bob. Bob will compute f (message), or "encryption function", on the ciphertext, without knowledge of the secret key, and then return the encrypted result, which only Alice is able to decrypt. Therefore, Bob will never be able to find out anything about the "message" (Hu, 2013). To overcome the limitations established previously, several methods of HE were proposed. Those approaches are shown in the following Table I.

The disadvantage of these approaches is that they only support one operator at the same time. For example, only XOR, addition or multiplication. However, Gentry (2009a) introduced the first FHE scheme, which is able to process both addition and multiplication operations in the encrypted domain at the same time. It allows one to use untrusted computing resources without the risk of revealing sensitive data. In addition, eventually, this scheme seems to be secure and therefore solves the problem mentioned in Rivest *et al.* (1978). FHE has three main variants:

(1) based on ideal lattices (Gentry, 2009b; Gentry and Halevi, 2011);

(2) based on integers (Coron *et al.*, 2011; Van Dijk *et al.*, 2010); and

(3) based on ring learning with errors (Brakerski *et al.*, 2012; Zvika, 2011).

Table I.
HE approaches

Year	Scheme	Homomorphism
1978	Textbook RSA	Multiplicative
1978	Textbook ElGamal	Multiplicative
1984	Goldwasser Micali	XOR
1994	Benaloh	Additive
1999	Paillier scheme	Additive
2000	Paillier ECC variations	Additive
1998	Naccache–Stern	Additive
2007	Kawachi–Tanaka–Xagawa	Additive
1998	Okamoto–Uchiyama	Additive
2005	Boneh–Goh–Nissim	2-DNF formulas
2010	Melchor–Gaborit–Herranz	doper-multiplication

Moreover, there is interest in the use of homomorphic schemes for a number of specific applications – for instance, cloud computing, electronic voting, data aggregation in distributed networks, biometrics and privacy-preserving data mining.

2.1 FHE construction

Basically, the FHE scheme can be used in both symmetric (secret key) and asymmetric (public key) ciphers. Consequently, there are four algorithms needed to establish the FHE scheme: key generation, encrypt, decrypt and evaluate (Gentry, 2009a). They are run in a time polynomial (λ) , with λ as a security parameter. In case of a symmetric cipher, λ is used in the key generation algorithm to create the key that will be used in the encrypt and decrypt algorithms, whereas in an asymmetric cipher, λ is used to create two keys: the public key *pk*, which will be available to anyone, and the secret decryption key *sk*. Commonly, the public key is used to encrypt a message, encrypt (*pk*, message), and the secret decryption key is used to decrypt a ciphertext, decrypt (*sk*, ciphertext). The evaluate algorithm evaluates the result of the computation f on the ciphertext C_1, C_2, $C_3 \ldots C_t$ using the public key *pk*, and the result is stored in a *Ciphertext C*. Finally, when *Ciphertext C* is decrypted using the decrypt algorithm, the output is f (m) , where m represents the plaintext.

3. Protocol

The biometric authentication protocol contains three parts:

(1) The key management part (1) is run on an authentication server (AS), which is used to set up the FHE parameters to generate public and secret keys. Such parameters are chosen based on Brakerski *et al.* (2012) and are set in our program as follows: security parameters (λ), set to 100; the number of levels of arithmetic circuit (L), set to 2; the degree of polynomials to be evaluated (d), set to 5; the message of the space modulus (t), set to 1,117; the noise distribution (X), set to 8; and a secure random number. As mentioned, the construction of this FHE scheme is based on Brakerski *et al.* (2012). Lastly, the FHE secret decryption key is stored on the AS, and the FHE public key is sent to the other parties involved in the authentication process, such as the client side (CS) and the database server (DBS). In this protocol, it is assumed that the AS and the CS are in the same organization (SITE A), and the DBS is in a different one (SITE B).

(2) In the enrolment part, a biometric template or a group of them are encrypted and stored in the database. Initially, in the CS, the biometric image is taken and processed according to the iris biometric (segmentation, normalization, and encoding), and it is converted into a binary vector of 2048 bits. Then, it is encrypted (with the FHE public key) bit-by-bit and sent to the DBS along with a given user ID for storage.

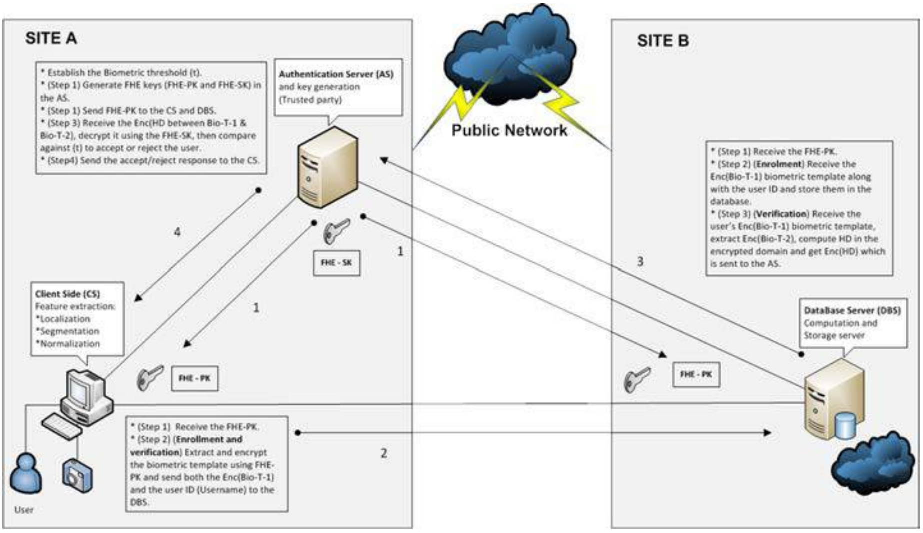
(3) The final part is authentication (3), where a given biometric template is encrypted and compared in the encrypted domain against the registered biometric template in the database in either the identification or verification processes. In this part, we consider two scenarios: (a) *Verification* – in this process, an individual is asked to provide his/her user ID, and his/her biometric image is taken, processed according to the iris biometric (segmentation, normalization and encoding) and then converted into a binary vector and encrypted (with the FHE public key) bit-by-bit on the CS. Then, both user ID and

the encrypted biometric template (Bio-T-1) are sent to the DBS. The DBS receives the package information and gets the enrolled template encrypted (Bio-T-2) according to the user ID received. In the encrypted domain, Bio-T-1 and Bio-T-2 templates are compared using the high definition (HD) in the encrypted domain and using the FHE public key. The encrypted result is sent to the AS. The AS decrypts the result using the FHE secret key and then compares it against the threshold to obtain the final decision to grant/deny access. (b) *Identification*– in this scenario, the same Bio-T-1 is taken, but in the DBS, the processes of HD computations and comparisons are done with all records ($N = Bio - T - k_{i=1,\, until\ i \leq N}$) in the DBS. The goal in this scenario is to find which HD result is below the threshold to identify the user ID and grant access or otherwise issue a message reporting that the user is not identified in the current database and is denied access.

The verification and identification protocols are illustrated in Figures 1 and 2, respectively. They describe the steps required and the parties involved in the protocols to grant or deny access.

4. Implementation

Our implementation consists of creating a biometric authentication system environment to test the effectiveness of the FHE scheme. To do so, we used a free Java library (under GNU license) named the Java Lattice Based Cryptography Library (jLBC) (De Caro, 2012). The FHE approach was tested with iris images from the BATH iris database (Sensors, 2009). We used Osiris version 4.1 (Guillaume *et al.*, 2012) to process the iris recognition process (segmentation, normalization, encoding and matching). MATLAB R2013R (MATLAB, 2013) was used to process the FRR versus the FAR to get the threshold, which is illustrated in Figure 3. The matching results (Hamming distance) to obtain the receiver operating characteristic can be seen in Figure 4, which shows a

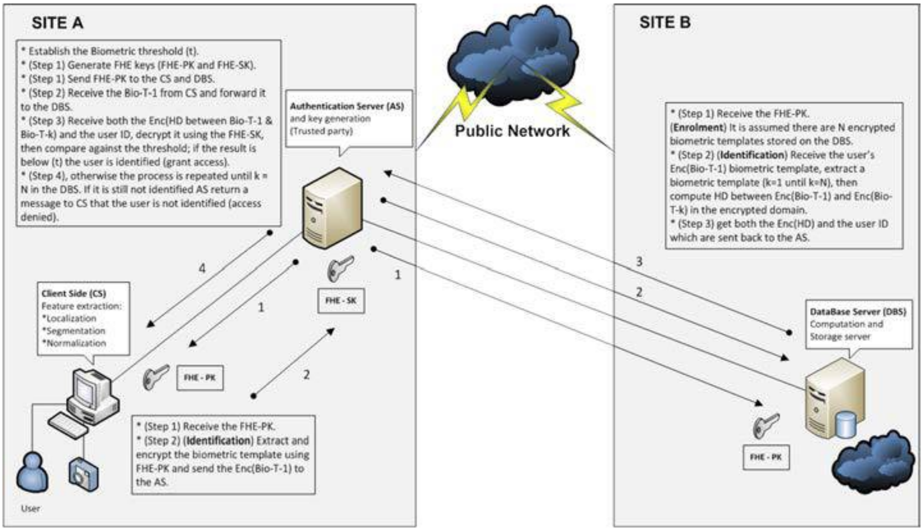


- * Establish the Biometric threshold (t).
- * (Step 1) Generate FHE keys (FHE-PK and FHE-SK) in the AS.
- * (Step 1) Send FHE-PK to the CS and DBS.
- * (Step 3) Receive the Enc(HD between Bio-T-1 & Bio-T-2), decrypt it using the FHE-SK, then compare against (t) to accept or reject the user.
- * (Step4) Send the accept/reject response to the CS.

- * (Step 1) Receive the FHE-PK.
- * (Step 2) (**Enrolment**) Receive the Enc(Bio-T-1) biometric template along with the user ID and store them in the database.
- * (Step 3) (**Verification**) Receive the user's Enc(Bio-T-1) biometric template, extract Enc(Bio-T-2), compute HD in the encrypted domain and get Enc(HD) which is sent to the AS.

- * (Step 1) Receive the FHE-PK.
- * (Step 2) (**Enrollment and verification**) Extract and encrypt the biometric template using FHE-PK and send both the Enc(Bio-T-1) and the user ID (Username) to the DBS.

Figure 1.
FHE implementation – authentication (verification)



* Establish the Biometric threshold (t).
* (Step 1) Generate FHE keys (FHE-PK and FHE-SK).
* (Step 1) Send FHE-PK to the CS and DBS.
* (Step 2) Receive the Bio-T-1 from CS and forward it to the DBS.
* (Step 3) Receive both the Enc(HD between Bio-T-1 & Bio-T-k) and the user ID, decrypt it using the FHE-SK, then compare against the threshold; if the result is below (t) the user is identified (grant access).
* (Step 4), otherwise the process is repeated until k = N in the DBS. If it is still not identified AS return a message to CS that the user is not identified (access denied).

* (Step 1) Receive the FHE-PK.
* (Step 2) (**Identification**) Extract and encrypt the biometric template using FHE-PK and send the Enc(Bio-T-1) to the AS.

* (Step 1) Receive the FHE-PK.
(**Enrolment**) It is assumed there are N encrypted biometric templates stored on the DBS.
* (Step 2) (**Identification**) Receive the user's Enc(Bio-T-1) biometric template, extract a biometric template (k=1 until k=N), then compute HD between Enc(Bio-T-1) and Enc(Bio-T-k) in the encrypted domain.
* (Step 3) get both the Enc(HD) and the user ID which are sent back to the AS.

Figure 2.
FHE implementation – authentication (identification)

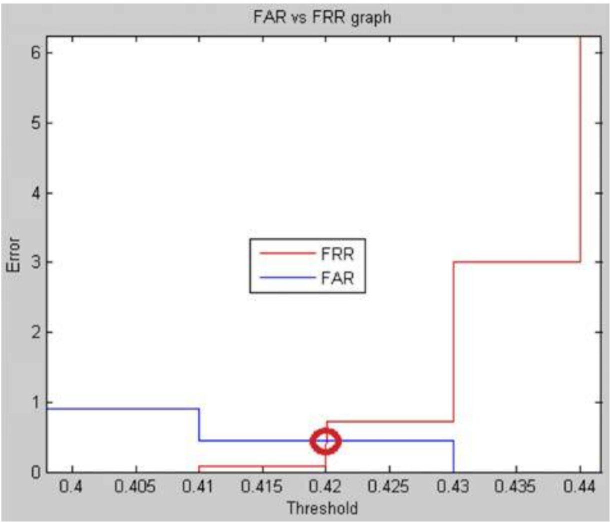


Figure 3.
FAR versus FRR; zoom to identify the threshold

high-accuracy performance for the iris sample. As a result, it is unlikely that an impostor will be accepted in the biometric authentication system. This can be also seen in the detection error trade-off in Figure 5. This experiment ran on an Intel Core i7-3630QM at 2.40 GHz with a 16GB memory, and the implementation was done in Java.

The threshold was set using the BATH iris database along with Osiris 4.1 and MATLAB. In total, 11 users from the BATH database were used, wherein each of them has 20 images. These 220 images were processed using Osiris 4.1 for segmentation, normalization and encoding. For the matching step, 1,100 samples for inter (imposter users) and 220 samples for intra (genuine users) were processed. By processing the

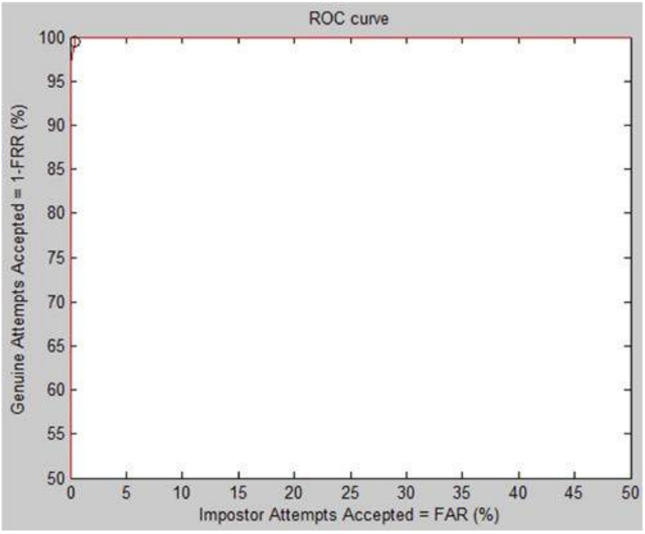


Figure 4.
ROC curve

matching results in MATLAB along with the scripts for the performance evaluation in Mayoue (2007), an energy efficiency ratio (EER) with a value of 0.5909 was obtained for this dataset, as illustrated in Figure 5. This result is demonstrated in Figure 3, which plots the decision threshold against FAR and FRR. According to the EER results, the threshold to 0.42 is also shown.

4.1 HD circuit

Commonly, the HD is used in the matching of the iris biometric process, and basically, it measures the number of bits, which are the same for two iris biometric templates. The result will recognize whether both templates are from the same iris (or same individual) or not. The HD can be calculated as $HD = \Sigma_{i=0}^{i=N} A_i(XOR)B_i / N$. Likewise, the function f in the FHE scheme can be represented as a Boolean circuit; therefore, the XOR $(x, y) = x + y - 2*x*y$. Therefore, this Boolean circuit is used in the implementation of biometric authentication systems using FHE. A genuine user is identified when the $HD < t$; also, under the assumption that $Z = \Sigma_{i=0}^{i=N} A_i(XOR)\ B_i$, we can say that $HD = Z/N \Rightarrow Z/N < t \Rightarrow Z < t*N$. For the chosen dataset of iris biometrics, the threshold is 0.42, and the length of the biometric vector is 2048, so we can conclude that the result of bit-by-bit XOR and the count of the number of bits “1” must be less than ≈ 860 to identify a biometric template as a genuine user; otherwise, it is an imposter. In this implementation, the result of z is compared to 860 to determine if the system is able to grant or deny access.

4.2 Results

In either the verification or identification processes, we are comparing two biometric templates to grant or deny access. As mentioned, each biometric template has a 2048-bit length, so it is running the Boolean circuit for HD computation in the encrypted domain 2,048 times. Consequently, it was found that the time required is around 10 minutes, as

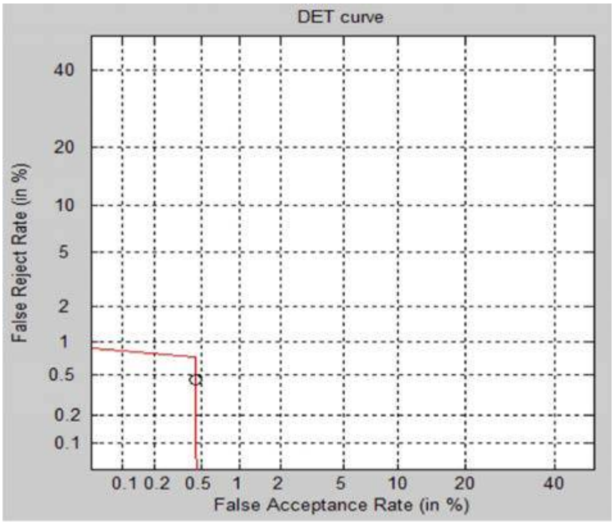


Figure 5. DET curve

illustrated in Table II. Similarly, the database size for 1 user and for 11 users (the whole iris biometric sample) is shown in Table III.

5. Evaluation

In this section, we evaluate the implementation of a biometric authentication system using the FHE scheme in terms of security, mainly some common points of attack a biometric system is vulnerable to, such as ideal biometric template protection properties and the FHE scheme properties. The implementation is also evaluated in terms of performance to consider how far this scheme can be used in real implementations.

5.1 Security analysis

To evaluate the privacy of using FHE in a biometric authentication system, the analysis was divided into three subtopics: analysis by points of attack, analysis by desirable properties and analysis by FHE properties.

FHE function	Time
Keygen	≈ 26,649 ms (26.649 seconds)
Encrypt	≈ 230,318 ms (3.8 minutes)
Hamming distance	≈ 415,820 ms (6.5 minutes)
Decrypt and comparison	≈ 49 ms (0.49 seconds)
Total	≈ 64,6187 ms (10.4 minutes)

Table II. Implementation: time for two biometric template comparisons

Database	Size
1 user	517 KB (0.504 MB)
11 users	5,676 KB (5.542 MB)

Table III. Implementation: database size

5.1.1 Analysis by points of attack. The points of attack in a biometric authentication systems are defined in Campisi, 2013; Alimi *et al.*, 2011; Ratha *et al.*, 2003; Jain *et al.*, 2008; Rathgeb and Uhl, 2011; Ratha *et al.*, 2001; Bolle *et al.*, 2002, which identified the points of attack as the sensor, feature extraction, matching, data storage, final decision point and all the channels that link them.

Among the attacks identified at the sensor point are:

The coercive attack, in which an authorized individual is violently forced to present his/her biometric to the sensor to get access to the authentication system.

Spoofing and mimicry attacks, where an attacker tries to masquerade as an authentic user or imitates the user's biometric features to get access.

- Device substitution attack is exploited when the physical device is modified or replaced for another device, which is used to capture the user's biometric features.
- The denial of service (DoS) attack is used to decrease the availability of the system.

As shown in Figure 6, on the client side entity, these attacks are hardware oriented, and the FHE scheme is unable to protect the user's biometric data when these attacks materialize. Likewise, the feature extraction point can be attacked either by inserting imposter data or replacing any hardware component, which results in the attacker forcing the authentication system to accept the fake biometric features, thereby getting into the system. The FHE scheme cannot be used against attacks at the sensor and feature extraction points because they are hardware oriented. As a countermeasure, one should apply best or recommended practices (Kindt, 2013) to reduce the risk probability of such attacks – for instance, by setting organization and security policies along with the international certification of a biometric system to preserve privacy against these attacks. Furthermore, the feature extraction has a software component. In our implementation, we used Osiris 4.1, and evidently, this software can be integrated with

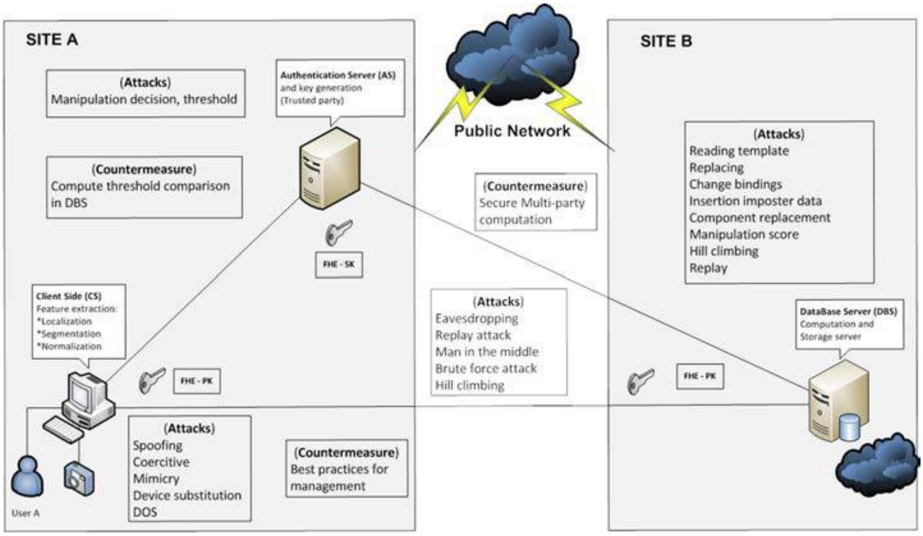


Figure 6.
Security analysis based on point of attacks

FHE implementation. This will increase the protection of privacy because it directly creates a ciphertext based on the iris input image.

It is important to consider regulations established by entities or governments as well as developing policies or guidelines on implementing privacy rules. Belguechi *et al.* (2011) and Kindt (2013) mention these considerations and also emphasize the idea of implementing any technology under the "Privacy by Design" (PbD) paradigm, which basically consists of integrating privacy and security protection at the initial phases of any development of information technologies. Moreover, privacy-enhancing technologies (PETs) have emerged to overcome the security and privacy limitations of biometric systems. They follow some foundation principles for any approach; for example, it should be impossible to recover the user's original biometric data, and it should use few of a user's personal data for any authentication process. As a complement, it is said that PETs require some improvements in its framework to provide better protection of biometric data (Scheirer and Boult, 2007).

As in every field of information technology, the guidelines or standards should be part of the implementation process to ensure the applicability of the correct protection approach for the biometric system. Kindt (2013) remarks that the ten best practices are based on the PbD paradigm, which contributes to the proper design, development and implementation of biometric system approaches in terms of their security and privacy, mainly in scenarios such as authentication and access control.

Attacks to the matching point are identified as:

- manipulation of the matching score, where the Hamming distance result is modified to affect the final decision;
- the replay attack, where previously captured biometric data are used;
- the attacker tries to modify the biometric data continuously in a hill-climbing attack until he gets a useful matching result to grant access; and
- the components, in either the hardware or software, of the matching point can be changed according to the attacker's needs.

The biometric template is encrypted using the FHE scheme after the feature extraction stage, and the HD is computed in the encrypted domain on the DBS in both authentication types: verification and identification. Besides that, the secret key (FHE-SK) is never placed in the DBS and is not even transmitted to any other party. This indicates that the biometric template is always protected against the above-mentioned attacks.

Moreover, the encrypted biometric templates are stored and are always kept encrypted in the database during the process of biometric authentication: enrolment, verification and identification. In either case, the database is vulnerable to reading or replacing biometric templates, implying that raw biometric data will not be revealed.

However, after the HD result is computed on the DBS, it needs to be compared against the threshold to determine if the user is genuine or an imposter. In the current implementation, this comparison is performed in the AS. This may open a vulnerability of privacy because the AS is somehow able to learn about the user's biometric template – for example, when the AS decrypts the cipher (which comes from the DBS), the number of bits, which is different between the two templates, is revealed. It is clear that this information is not the original or raw template, but this information can be useful to learn about the raw biometric template as presented in Figure 6. This vulnerability can be referred to as hill climbing. As a countermeasure, the threshold comparison should be computed in the encrypted domain; as a result, this will get two possible outputs (yes/no) when the AS decrypts the ciphertext. This will increase the level of privacy and will prevent an attacker learning about the raw user's biometric template from the AS.

Lastly, the attacks that are presented in the channels are the following:

- Eavesdropping when the channels are sniffed for biometric data.
- Man in the middle attacks when the communication between two points are intercepted, and these points are unaware that the communication is compromised. When the attacker tries to get into the system by presenting varying biometric templates, the attack is called brute force.
- The manipulation of the decision attack in which the final decision results are captured and modified.
- The attacks replay, hill climbing and the manipulation of the matching score, which are mentioned and explained above, are also present in the channels.

As already mentioned, the FHE scheme can offer a high level of privacy because all data going through the channels are always encrypted, and the private or secret decryption key is never sent to another party; in fact, the calculations for getting the Hamming distance are done in the encrypted domain using the parties' public keys.

However, it is assumed that our implementation is based on honest management (or a semi-honest model); nevertheless, this implementation has to consider a dishonest management (or malicious model) to provide a high level of privacy protection. This can be done by designing the implementation with the perfect fairness or ideal model to avoid scenarios such as when a party refuses to participate in the protocol, when local inputs are substituted by the parties or when a party cancels the protocol prematurely (Goldreich, 2009). In addition, the implementation of this concept will carry out a randomness component in the encrypted domain; consequently, the time required to perform this computation will also increase considerably.

5.1.2 Analysis by desirable properties. Another point of view where privacy-preserving data can be evaluated by analyzing the ideal properties was proposed initially in Alimi *et al.*, 2011; Campisi, 2013; Jain *et al.*, 2008; and Ratha *et al.*, 2003. These properties are non-reversibility, accuracy, cancellability, diversity, revocability, performance and randomness.

Non-reversibility is the inability to reconstruct a user's original or raw template from an encrypted one. The FHE scheme uses semantic security (Gentry, 2010), which is mentioned as a strong security property based on probabilistic, meaning that it is computationally infeasible to extract any plaintext information from the ciphertext (Goldwasser and Micali, 1984). This property increases the difficulty of getting to the plaintext (original template) if an attacker knows the length of the message M and uses any computational function polynomial A with FHE-PK (public key) and the ciphertext (encrypted biometric template). This property can be explained when the HD computations are performed on DBS with the FHE-PK, and the ciphertext is modified homomorphically without the need of the FHE-SK. In fact, as mentioned, the FHE-SK stored in the AS is never sent to the DBS.

Accuracy and *performance* preserve the accuracy of recognition when the biometric template is processed. As previously mentioned, the FHE scheme's evaluate algorithm allows the performance of arbitrary computation in the encrypted domain; the noise is handled using a refresh function to keep the protected information accurate after processing any function or circuits. In our implementation, HD computation is performed in the encrypted domain at the DBS, and the results we get after decryption at the AS are the results expected when compared with the entire operation being performed on the biometric template directly. This substantiates how the FHE scheme maintains privacy without compromising biometric recognition, which is also accurate. Therefore, we can rely on the result of biometric performance (FAR and FRR) for the biometric samples taken from the iris database (Sensors, 2009).

Renewability is the ability to create many encrypted templates based on the same original template. As explained in the non-reversibility property, the FHE scheme uses semantic security, and this enables the creation of many encrypted templates with the same original template.

Diversity is the difficulty of cross-matching encrypted templates in different applications. By using the FHE scheme, all the biometric templates are traveling in an encrypted form, computing the HD in the encrypted domain, and the final decision is only decrypted by the party who has the private key. This private key is never sent to any other party during the process of enrolment or authentication. As a result, the FHE scheme prevents cross-matching, as each application must have its own key management infrastructure.

Revocability is defined as an easy way to revoke a compromised biometric template and recreate a new one based on the same raw biometric template, and *randomness* implies that the knowledge of multiple revoked templates does not help to predict a following accepted one. We started with the premise that the biometric data are always protected because of both the FHE scheme's properties and the biometric system using the FHE scheme. In fact, the main three algorithms for the FHE scheme (key generation, encryption and decryption) use a randomized algorithm (Yasuda *et al.*, 2013c; Gentry, 2010) that meets both desirable properties: revocability and randomness.

5.1.3 Analysis by FHE properties. Now we shall examine how the FHE scheme can be attacked. The FHE scheme used in this work is based on ideal lattices (Gentry, 2009b). The lattice-reduction algorithm is used to attack the FHE scheme, as is shown and demonstrated in Yasuda *et al.*, 2013c; Nguyen, 2011; and Chen and Nguyen, 2011. The objective of this algorithm is to find the short and near orthogonal vectors among the two types (exact and approximation). The most suitable type to test the FHE scheme is the approximation algorithm, which is faster than the exact approach. The most famous approximation algorithm is Lenstra, Lenstra and Lovász (LLL) (Lenstra *et al.* (1982). Moreover, based on the FHE scheme challenges published by Gentry and Halevi (2010), these were named as follows: toy, small, medium and large, which are related to the Hermite constant. It was revealed that as the Hermite constant increases, the lattice problems become easier to solve. The Hermite constant given for the challenges are 1.67, 1.14, 1.03 and 1.0081, respectively (Chen and Nguyen, 2011).

It was found that an attack to the toy, small and medium is not difficult to solve. In fact, using the LLL, with lattice dimensions of 512, 2,048 and 8,192, the estimated time to decrypt a ciphertext is "30 core days", "more than 45 core years" and "more than 68,582 core years", respectively. As a result, it is recommended that a minimum lattice dimension of 10,000 be used to make the FHE scheme secure (Yasuda *et al.*, 2013c; Chen and Nguyen, 2011). In our implementation, this variable is represented by $n = n(\lambda)$, where λ is the security parameter, which is used to represent 2^{λ} as the security against common lattice attacks, such as lattice-reduction attack (Brakerski *et al.*, 2012); setting the value of security parameter to "100" is also suggested.

5.2 Performance analysis

As previously shown in Section 4.2, the process of matching two biometric templates in the encrypted domain takes around 10.4 minutes. This process consists of FHE key generation, FHE encryption, Hamming distance computation, FHE decryption and threshold comparison to obtain the results (granted or denied). It also applies to either verification or identification. A similar implementation was done in Pötzelsberger (2013) where a biometric identification process was implemented by using three versions of partially HE: Goldwasser–Micali, Paillier and Paillier Chunkwise.

In that implementation, an iris sample was used as the biometric trait, which was encrypted bit-by-bit using a Paillier HE algorithm. Besides that, to get the HD in the matching process, the XOR circuit in the encrypted domain along with a permutation function was performed to preserve privacy. The remainder computations for the Hamming distance were performed in plaintext. Under these conditions, the process of identification for a user takes around 9.2 minutes; however, it does not mention if this time includes key generation, encryption, decryption, HD and comparison, as is explained in the present work. In our implementation, we included the time for the whole process of matching two biometric templates. In both implementations, the processing time for two biometric templates without encryption is less than one second.

Although Paillier HE seems like a better approach in comparison to the FHE scheme in terms of time performance, it may not offer effective protection because of its limitations when computing in the encrypted domain. As mentioned, Paillier HE performs just an XOR in the encrypted domain, with a permutation of the 2048 bits of a biometric template. Clearly, this approach is not enough to preserve the privacy of this kind of information, as this scheme is vulnerable to the hill-climbing attack, which will enable an adversary to obtain raw biometric data. Moreover, the Paillier scheme itself has limitations in its scheme; it just allows performing additions in the encrypted domain, and possibly, this explains why the author performs only XOR in his implementation.

On the other hand, our FHE scheme implementation went further; we performed Hamming distance computation in the encrypted domain, which included both additions and multiplications, but this implementation still required improvements to preserve privacy. For example, we found the need to include the process of comparison of the Hamming distance in the encrypted domain; this will yield two options (yes/no or accept/reject), which reduces the privacy risk significantly. Nevertheless, if this comparison computation is included, the time required to process a match between two templates will rise significantly. This explains why many researches, such as Bringer *et al.*, 2013; and Yasuda *et al.*, 2013b, state that the FHE scheme is impractical for real implementations.

There are other factors in the hardware and software that can influence the performance result. The FHE scheme was implemented in the Java language using a cryptography library also in Java. However, Java is a high-level language, and instead, low-level languages such as C or C++ will be efficient (Prechelt, 1999). Furthermore, a packed method can be used to improve the time performance and decrease the size of the database because it reduces the amount of computations in the encrypted domain. However, this approach is not applied in the implemented FHE scheme. Instead, it is applied in the somewhat HE scheme (Yasuda *et al.*, 2013a), where an impressive improvement is observed, particularly in time. It takes around 19 milliseconds, and this might be the target to use for the FHE scheme in real implementations. The other factor is hardware. In the FHE scheme, the implementation used desktop hardware architecture; nonetheless, performance can be improved just by using server hardware architecture.

6. Conclusion and future work

This paper concludes that the FHE approach is effective and can provide protection for preserving the privacy of biometric data because it provides unlimited computation in the encrypted domain, and the private key (FHE-SK) is not sent to any party during the process of the authentication protocol. However, some future potential improvements must be made to the current implementation. For example:

- use the recommended practices and integrate the FHE scheme with Osiris 4.1 (software of feature transformation) to reduce the risk of hardware-oriented attacks;
- besides the Hamming distance computation, compute the threshold comparison in the encrypted domain, as this will raise the privacy protection of the biometric template because it retrieves information, such as yes/no or 1/0, which does not have any relationship with raw biometric data;
- design the implementation based on the malicious model (dishonest management) to secure the protocol; and
- set the parameters of the FHE scheme appropriately to guard against common attacks.

It was demonstrated that the biggest limitations of the FHE approach are the performance and the size of the ciphertext. It was observed that there is a trade-off between computations in the encrypted domain and the time required for execution along with the size of the database. Future improvements can be made to deal with these limitations. For instance, implement the FHE scheme in a low-level language such as C or C++ to enhance the efficiency of the implementation, which was done in Java; and implement the parallelism FHE property to reduce the amount of computation in the encrypted domain. This property is also known as packing (Brakerski *et al.*, 2012), which was implemented using the somewhat HE scheme (Yasuda *et al.*, 2013a).

These improvements may open the possibility of real-time implementations not only in biometric systems but also in other fields, such as cloud computing, e-voting and data mining.

References

Alimi, V., Belguechi, R., Cherrier, E., Lacharme, P. and Rosenberger, C. (2011), "An overview on privacy preserving biometrics", in Yang, J. and Poh, N. (Eds), *Recent Application in Biometrics*, InTech, France, pp. 978-953.

Belguechi, R., Alimi, V., Cherrier, E., Patrick, L. and Rosenberger, C. (2011), *Recent Application in Biometrics*, in Yang, J. and Poh, N. (Eds), InTech, Rijeka.

Bolle, R.M., Connell, J.H. and Ratha, N.K. (2002), "Biometric perils and patches", *Pattern Recognition*, Vol. 35 No. 12, pp. 2727-2738.

Brakerski, Z., Gentry, C. and Vaikuntanathan, V. (2012), "(Leveled) fully homomorphic encryption without bootstrapping", *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference, ACM, Tokyo*.

Bringer, J., Chabanne, H. and Kindarji, B. (2011), "Identification with encrypted biometric data", *Security and Communication Networks*, Vol. 4 No. 5, pp. 548-562.

Bringer, J., Chabanne, H. and Patey, A. (2013), "Privacy-preserving biometric identification using secure multiparty computation: an overview and recent trends", *Signal Processing Magazine*, IEEE, Vol. 30 No. 2, pp. 42-52.

Caloyannides, M. (2003), "Privacy vs information technology", *Security & Privacy*, IEEE, Vol. 1 No. 1, pp. 100-103.

Campisi, P. (2013), "Security and privacy in biometrics: towards a holistic approach", *Security and Privacy in Biometrics*, Springer, Heidelberg, pp. 1-23.

Chen, Y. and Nguyen, P. (2011), "BKZ 2.0: better lattice security estimates", *Advances in Cryptology – ASIACRYPT 2011*, Vol. 7073 No. 1, pp. 1-20.

Coron, J.S., Mandal, A., Naccache, D. and Tibouchi, M. (2011), "Fully homomorphic encryption over the integers with shorter public keys", *Advances in Cryptology – CRYPTO 2011*, Vol. 6841 No. 1, pp. 487-504.

Daugman, J. (1997), "Neural image processing strategies applied in real-time pattern recognition", *Real-Time Imaging*, Vol. 3 No. 3, pp. 157-171.

Daugman, J. (2000), *Biometric Decision Landscapes*, University of Cambridge, Computer Laboratory, Cambridge.

De Caro, A. (2012), "Java lattice based cryptography library", available at: http://gas.dia.unisa.it/projects/jlbc/index.html (accessed 15 May 2014).

Gentry, C. (2009a), "A fully homomorphic encryption scheme", Thesis, Stanford University, CA.

Gentry, C. (2009b), "Fully homomorphic encryption using ideal lattices", *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, ACM, Tokyo*.

Gentry, C. (2010), "Computing arbitrary functions of encrypted data", *Communications of the ACM*, Vol. 53 No. 3, pp. 97-105.

Gentry, C. and Halevi, S. (2010), "Public challenges for fully-homomorphic encryption", available at: http://researcher.ibm.com/researcher/view_project.php?id=1548 (accessed 20 June 2014).

Gentry, C. and Halevi, S. (2011), "Implementing gentry's fully-homomorphic encryption scheme", *Advances in Cryptology – EUROCRYPT 2011*, Vol. 6632 No. 1, pp. 129-148.

Goh, A. and Ngo, D.C. (2003), "Computation of cryptographic keys from face biometrics", *Communications and Multimedia Security: Advanced Techniques for Network and Data Protection*, Torino, pp. 1-13.

Goldreich, O. (2009), *Foundations of Cryptography: Volume 2, Basic Applications*, Cambridge University Press, Cambridge.

Goldwasser, S. and Micali, S. (1984), "Probabilistic encryption", *Journal of Computer and System Sciences*, Vol. 28 No. 2, pp. 270-299.

Guillaume, S., Bernadette, D., Sonia, G.-S. and Nadia, O. (2012), "A biometric reference system for iris OSIRIS version 4.1", available at: http://svnext.it-sudparis.eu/svnview2-eph/ref_syst//Iris_Osiris_v4.1/doc/Documentation_OSIRIS_v4.1_.pdf (accessed 20 April 2014).

Hu, Y. (2013), "Improving the efficiency of homomorphic encryption schemes", Doctoral dissertation thesis, KIIT University, India.

Jain, A.K., Ross, A. and Prabhakar, S. (2004), "An introduction to biometric recognition", *Circuits and Systems for Video Technology, IEEE Transactions on*, Vol. 14 No. 1, pp. 4-20.

Jain, A.K., Nandakumar, K. and Nagar, A. (2008), "Biometric template security", *EURASIP Journal of Advanced Signal Process*, Vol. 2008 No. 113, pp. 1-17.

Juels, A. and Sudan, M. (2006), "A fuzzy vault scheme", *Designs, Codes and Cryptography*, Vol. 38 No. 2, pp. 237-257.

Juels, A. and Wattenberg, M. (1999), "A fuzzy commitment scheme", *Proceedings of the 6th ACM Conference on Computer and Communications Security*, *ACM*, *New York, NY*.

Kindt, E. (2013), "Best practices for privacy and data protection for the processing of biometric data", *Security and Privacy in Biometrics*, Springer, London, pp. 339-367.

Lenstra, A.K., Lenstra, H.W. and Lovász, L. (1982), "Factoring polynomials with rational coefficients", *Mathematische Annalen*, Vol. 261 No. 4, pp. 515-534.

MATLAB (2013), *R2013R*, The MathWorks, Natick, MA.

Mayoue, A. (2007), "Performance evaluation of a biometric verification system", available at: http://svnext.it-sudparis.eu/svnview2-eph/ref_syst/Tools/PerformanceEvaluation/doc/ (accessed 10 April 2014).

Mollin, R.A. (2006), *An Introduction to Cryptography*, CRC Press, Taylor & Francis.

Nguyen, P. (2011), "Lattice reduction algorithms: theory and practice", *Advances in Cryptology – EUROCRYPT 2011*, Vol. 6632 No. 1, pp. 2-6.

Pötzelsberger, G. (2013), "KV Web security: applications of homomorphic encryption", available at: www.fim.uni-linz.ac.at/lva/Web_Security/Abgaben/Poetzelsberger-Homomorphic.pdf (accessed 20 March 2014).

Prabhakar, S., Pankanti, S. and Jain, A.K. (2003), "Biometric recognition: security and privacy concerns", *Security & Privacy*, IEEE, Vol. 1 No. 2, pp. 33-42.

Prechelt, L. (1999), "Comparing Java vs C/C++ efficiency differences to interpersonal differences", *Communications ACM*, Vol. 42 No. 10, pp. 109-112.

Ratha, N.K., Connell, J.H. and Bolle, R.M. (2001), "Enhancing security and privacy in biometrics-based authentication systems", *IBM Systems Journal*, available from: ProQuest Central, Vol. 40 No. 3, pp. 614-634.

Ratha, N.K., Connell, J.H. and Bolle, R.M. (2003), "Biometrics break-ins and band-aids", *Pattern Recognition Letters*, Vol. 24 No. 13, pp. 2105-2113.

Rathgeb, C. and Uhl, A. (2011), "A survey on biometric cryptosystems and cancelable biometrics", *EURASIP Journal on Information Security*, Vol. 2011 No. 1, pp. 1-25.

Rivest, R.L., Adleman, L. and Dertouzos, M.L. (1978), "On data banks and privacy homomorphisms", *Foundations of secure computation*, Vol. 32 No. 4, pp. 169-178.

Scheirer, W.J. and Boult, T.E. (2007), "Cracking fuzzy vaults and biometric encryption", *Biometrics Symposium*, Baltimore, MD, pp. 1-6.

Sensors, S. (2009), "IRISBASE – A resource for research and evaluation", available at: www.smartsensors.co.uk/irisweb/ (accessed 3 April 2014).

Simoens, K., Tuyls, P. and Preneel, B. (2009), "Privacy weaknesses in biometric sketches", *Security and Privacy, 2009 30th IEEE Symposium*, *New York, NY*, pp. 188-203.

Škorić, B. (2010), "Security with noisy data", *Information Hiding*, Springer, Calgary.

Stallings, W. and Brown, L. (2012), *Computer Security: Principles and Practice*, Pearson, Boston.

Van Dijk, M., Gentry, C., Halevi, S. and Vaikuntanathan, V. (2010), "Fully homomorphic encryption over the integers", *Advances in Cryptology – EUROCRYPT*, Springer, Heidelberg, pp. 24-43.

Yasuda, M., Shimoyama, T., Kogure, J., Yokoyama, K. and Koshiba, T. (2013a), "Packed homomorphic encryption based on ideal lattices and its application to biometrics", in Cuzzocrea, A., Kittl, C., Simos, D., Weippl, E. and Xu, L. (Eds), *Security Engineering and Intelligence Informatics*, Springer, Berlin Heidelberg, Vol. 8128, pp. 55-74.

Yasuda, M., Shimoyama, T., Kogure, J., Yokoyama, K. and Koshiba, T. (2013b), "Secure pattern matching using somewhat homomorphic encryption", *Proceedings of the 2013 ACM Workshop on Cloud Computing Security Workshop*, *New York, NY*, pp. 65-76.

Yasuda, M., Yajima, J., Shimoyama, T. and Kogure, J. (2013c), "Analysis of lattice reduction attack against the somewhat homomorphic encryption based on ideal lattices", *Public Key Infrastructures, Services and Applications*, Vol. 7868 No. 1, pp. 1-16.

Ye, S., Luo, Y., Zhao, J. and Cheung, S.C.S. (2009), "Anonymous biometric access control", *EURASIP Journal of Information Security*, Vol. 2009 No. 1, pp. 1-17.

Zvika, B. (2011), "Efficient fully homomorphic encryption from (standard) LWE", *IEEE 54th Annual Symposium on Foundations of Computer Science*, *Berkeley*, pp. 97-106.

About the authors

Wilson Abel Alberto Torres (MCSA, MCTS, MCITP and MCP) is an active researcher with more than 10 years' experience working in the information technology (IT) industry. He received his Bachelor's degree in Computer Systems Engineering from Los Libertadores University, Colombia, and he holds a Master's (Honours) degree from Monash University, Australia. His expertise and research interests are networking, IT security and privacy. He is currently a Teaching Associate at the Faculty of Information Technology, Monash University, Australia.

Dr Nandita Bhattacharjee is a Senior Academician at the Faculty of Information Technology, Monash University, Australia. She has more than 30 years' experience in research and development in academia and industry. Her research interests are in the areas of signal/image processing, advanced digital design, biometric authentication and identification for information security, including privacy. She holds a Bachelor's and a Master's degree in Electronics and Telecommunications from the Indian Institute of Technology, Bombay, India, and a PhD in Computer Science from Monash University, Australia. Nandita Bhattacharjee is the corresponding author and can be contacted at: nandita.bhattacharjee@monash.edu

Bala Srinivasan is a Professor of Information Technology at the Faculty of Information Technology, Monash University, Australia. He has more than 30 years' experience in academia, industry and research organizations. He has authored and jointly edited technical books and has published articles in international journals and conferences in the areas of multimedia databases, data communications and data mining and distributed systems. He has a Bachelor of Engineering (Honours) degree in Electronics and Communication Engineering from Guindy Engineering College, University of Madras, India (receiving a gold medal), and he received Master's and PhD degrees in Computer Science from the Indian Institute of Technology, Kanpur, India.