Al-Assam, H., Hassan, W. and Zeadally, S., 2019. Automated Biometric Authentication with Cloud Computing. In Biometric-Based Physical and Cybersecurity Systems (pp. 455-475). Springer, Cham.

Automated Biometric Authentication with Cloud Computing

Hisham Al-Assam, Waleed Hassan

Applied Computing Department, University of Buckingham, Buckingham, MK18 1EG, UK

Sherali Zeadally

College of Communication and Information, University of Kentucky, Lexington, KY, 40506, USA

Abstract

The convenience provided by cloud computing has led to an increasing trend of many business organizations, government agencies and individual customers to migrate their services and data into cloud environments. However, once clients' data is migrated, the overall security control will be immedicably shifted form data owners to the hand of cloud service providers. In fact, most cloud clients do not even know where their data is physically stored, and therefore the question of how to limit data access to authorized users has been one of the biggest challenges in cloud environments. Although security tokens and passwords are widely used form of remote user authentication, they can be lost or stolen as they are not linked with the identity of data owner. Therefore, biometric based authentication can potentially offer a practical and reliable option for remote access control. This chapter starts with a brief introduction that covers the fundamental concepts of cloud computing and biometric based authentication. It then provides and in-depth discussions on authentication challenges for the cloud computing environment and the limitation of traditional solutions. This leads to the key sections related to biometric solutions for cloud computing in which we present state-of-art approaches that offer convenient and privacy-preserving authentication needed for cloud environment. The chapter argues that addressing privacy concerns surrounding the use of biometrics in cloud computing is one of the key challenges that has to be an integral part of any viable solution for any biometric-based authentication. It also argues that assuring cloud clients that their biometric templates will not be used without their permission to, for example, track them is not enough. Such solutions should make it technically infeasible to do so even if a cloud service provider wants to. This chapter explains a number of interesting solutions that have been recently proposed to improve security and, at the same time, maintain user privacy. Finally, we identify some challenges that still need to be addressed and highlight relevant Research Directions.

Keywords: Could computing, remote biometric authentication, fuzzy identify based encryption, multi-factor authentication.

Al-Assam, H., Hassan, W. and Zeadally, S., 2019. Automated Biometric Authentication with Cloud Computing. In Biometric-Based Physical and Cybersecurity Systems (pp. 455-475). Springer, Cham.

1. Introduction

Over the last few years, cloud computing has become one of the fastest growing IT environments for providing services to individuals and businesses of all sizes. Cloud computing, as defined by the National Institute of Standards and Technology (NIST), is "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" [1]. The so-called Cloud Service Providers (CSPs) are the key players in cloud computing responsible for providing clients with a wide range of services that vary from applications such Microsoft Office 365 and Google Docs to a complete infrastructure such as Amazon Elastic Compute Cloud (EC2) [2]. This introductory section provides the reader with a brief background on four related topics: 1) the main characteristics of cloud computing, delivery models, and deployment models, 2) security challenges in cloud computing, 3) biometric based recognition and, 4) the limitations of conventional biometric solutions for remote cloud authentication.

1.1. Cloud computing

The convenience provided by cloud computing has led to an increasing trend of many business organizations, government agencies and customers to migrate their services and data into cloud environments. The recent success of cloud computing in attracting such a great attention can be attributed to the following five characteristics [2]:

- *On-demand self-service*: A client can immediately get computing resources (e.g., CPU time, applications, network storage, etc.) without a need for human intervention at the CSP side.p
- *Broad network access*: Cloud resources are network accessible from different clients' applications installed on different platforms such as smart phones, tablets, PCs, and laptops.
- *Resource pooling*: The CSPs aggregate their resources to meet clients' need by utilizing multi-tenant approaches based on physical as well as virtual resources which can be dynamically added or withdrawn based on clients' requirements. The *pooling* factor means that the clients do not need to know where the resources are coming from or where the data is physically stored.
- *Rapid elasticity*: The capabilities of cloud services should be flexible enough to rapidly shrink or expand to meet the requirements of different clients at different times.
- *Measured* service: CSPs have the ability to measure any resources used by each tenant (client) using charge-per-use mechanisms.

Cloud services are typically delivered to clients using pre-packaged combinations of IT resources provided by CSPs based on one of the following three common cloud service models [3].

Software as a Service (SaaS): This model of delivery is also called "on-demand software". The software and associated data are centrally hosted on CSP's servers (i.e. instead of using a software install on Clients'

machine, they can use it as a service where no maintenance or upgrades are required). In this model, clients have no control or management permission over the underlying cloud infrastructure. Common examples of SaaS include Google Docs, Dropbox, and Microsoft Office 365.

Platform as a Service (PaaS): This kind of service is typically used by application developers. This type of service provides access to computing platforms that include operating systems, programming languages, software tools, databases, web servers, etc. The clients have control only over their deployed applications. Some examples of PaaS include Google AppEngine, Microsoft Azure, and Apache Stratos.

Infrastructure as a Service (IaaS): This delivery model supplies clients with computing resources (physical or more often virtual) processors, storage, firewalls, load balancers, virtual local area networks, and so on. Therefore, the clients are not only able to deploy and execute various software but they also have control over the operating systems, storage, processing power, and networking components. Amazon's EC2 is a very good example of IaaS.

The above three categories of services can be deployed in different environments. Deployment models define ownership and the size of cloud resources, and most importantly define who can access them. Currently, four basic models of deployment have been identified by the cloud community [3].

- **Private Cloud Computing**: The cloud infrastructure and services are offered exclusively to one enterprise, and it might be owned, managed as well as operated by the enterprise, a third party or a combination of both. This deployment model not only gets an optimal use of existing in-house resources but it also provides better data security and privacy. It should be noted that the cloud environment in this model might be located in or outside of the premises of the enterprise.
- **Community cloud computing**: The cloud infrastructure is shared by a group of clients or organizations to provide shared policies, values, and security procedures. The ownership, management, and operation of this model are given to one or more members of the group.
- **Public Cloud Computing**: The cloud infrastructure is open for public use. The ownership and management are given to business, academic institutes, government bodies, and so on.
- **Hybrid Cloud Computing**: More than one deployment models can be combined to form a hybrid cloud environment to meet clients' needs.

It can be argued that each type of service and deployment model meets the demands of some business more than others. For example, while a large enterprise might benefit from the private cloud, smaller businesses will most likely opt for a public cloud for cost consideration. Figure 1 illustrates typical cloud computing service layers along with their cost and timeline impact.

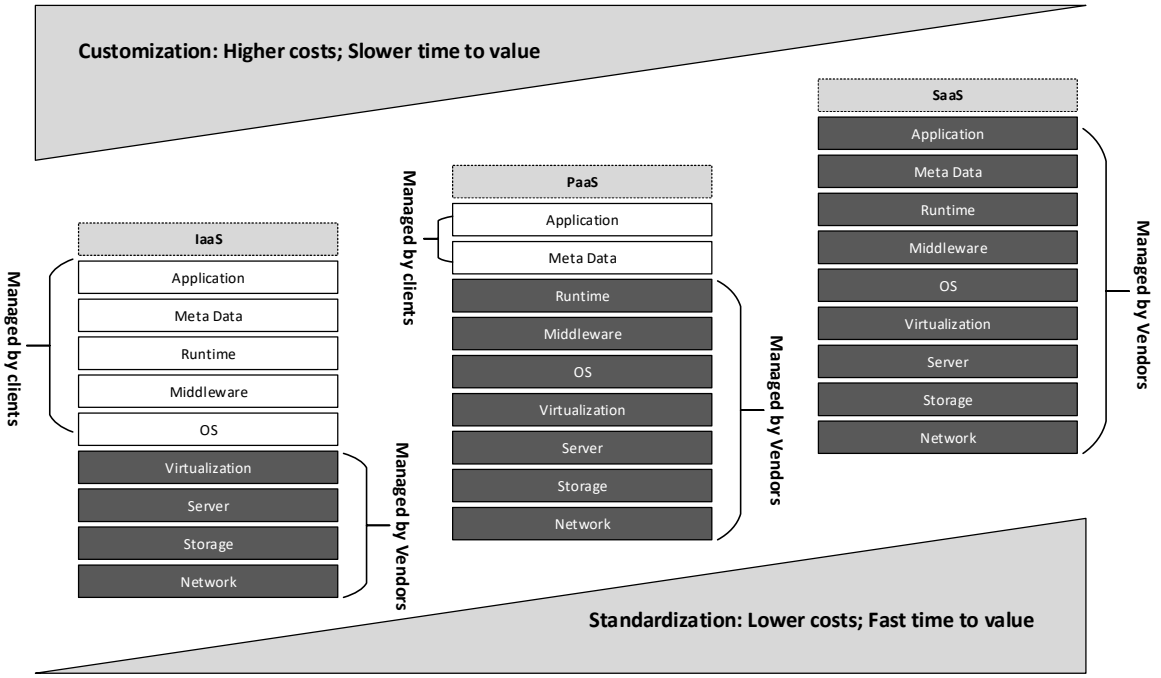

Figure 1. Typical Layers of Cloud Computing Services (adapted from [2])

1.2. Security challenges in cloud computing

Although cloud computing offers great advantages over other traditional IT solution, it poses serious security concerns. In fact, security and privacy are essential factors for an enterprise when deciding on whether to migrate their data, applications, and other relevant services to cloud environments. Service agreements between clients and CSPs tend to include details on how to access and utilize cloud services, service duration, and data storage and management when the contract ends [1]. However, the main challenge is how to guarantee that the data is accessible by authorized users only. When data owners decide to use the cloud environment, they rely entirely on third parties to make decisions about their data. Therefore, it is very important for data owners to have the right technologies or methods to prevent CSPs from utilizing such data without their permission. Both technical and non-technical methods have to provide effective means to fulfil this goal [4] [5]. A wide range of possible solutions have been proposed to implement different mechanisms to prevent unauthorized access to cloud data even by untrusted CSPs [6] [7] [8] [9] [10] [11]. In general, to address clients' concerns about security and privacy of the cloud environment, the following three essential challenges must be addressed [4]:

- **Outsourcing**: In the traditional IT environment, clients can exercise full control over their data. However, they usually lose all means of physical control over the data once it is migrated to cloud environments, which is the key security concern. To overcome this problem, clients need to ensure that the cloud services providers are trustworthy and are capable of meeting the requirements related to

secure data storage, correctness and integrity of cloud data and computation at all times, and maintaining clients' privacy.

- **Multi-tenancy**: Cloud environments can share their resources and services among multiple clients simultaneously. Both the virtual machines provided by CSPs and the cloud data of different clients are eventually located on a single physical machine based on particular resource allocation policy. Hence, a legitimate cloud client can potentially act as an adversary by exploiting some holes in the policies to gain unauthorized access to the data of other users.
- **Big data and intensive computation**: Cloud environment requires dealing with large volumes of data supported by powerful processing capabilities. Hence, traditional security techniques might be difficult to apply on such data because of the volume of high computation and communication overheads. For instance, to guarantee the integrity of remotely stored data, it is computationally infeasible to hash the whole data. Consequently, new strategies and protocols are needed to overcome such difficulties.

1.3. Biometric Authentication

In the cloud environment, a reliable identity management system is a key component to prevent identity theft and control access to different resources. Establishing the correct identity of a person is an essential task in any identity management system. Typically, there are three ways to authenticate an individual, each of which has its own advantages and limitations [12]:

- **Knowledge-based authentication**, or "something you know", that typically relies on a memorized password or PIN. A random 12-character password, for example, offers a strong security mechanism for user authentication. However, in practice, humans have difficulties in memorizing complex passwords, and passwords that they can easily remember are often short and therefore simple to guess or determined by a brute-force / dictionary attack.
- **Object-based authentication**, or "something you have", which relies on the physical possession of an object, such as a token. The main drawback of a physical token is that, when lost or stolen, an impostor can easily gain unauthorized access.
- **Identity-based authentication**, or "something you are", biometric-based authentication offers an advantage over other authentication factors in that a genuine user does not need to remember or carry anything. Moreover, biometric-based authentication is known to be more reliable than traditional authentication because it is directly linked with the identity of individuals. This is practically important for cloud environments as it associates data access with its ownership. However, biometric systems were not initially design for remote authentication in cloud environments. In fact, they can be subject to replay attack and, unlike other credentials such as PINs, passwords, or smart cards, once biometric related information is compromised, it is impossible to be changed again.

Biometric systems in general aim to identify or verify an individual's identity based on physical characteristics (e.g., face, iris, fingerprint, DNA, or hand geometry), and/or behavioural characteristics (e.g. speech, gait, or signature). A typical biometric system has two stages, enrolment and recognition. Figure 2

illustrates the process of the biometric enrolment stage, in which a user starts by presenting their biometric data to a biometric sensor (usually in a controlled environment). If the quality of the captured biometric sample is found to be adequate, the enrolment process proceeds to a pre-processing procedure to prepare the sample for the next step. A feature extraction technique is then used to extract a digital discriminating feature vector of the individual, called Biometric Template (BT), which will then be stored (often also called "enrolled") alongside the individual's identifier (ID) in a database.

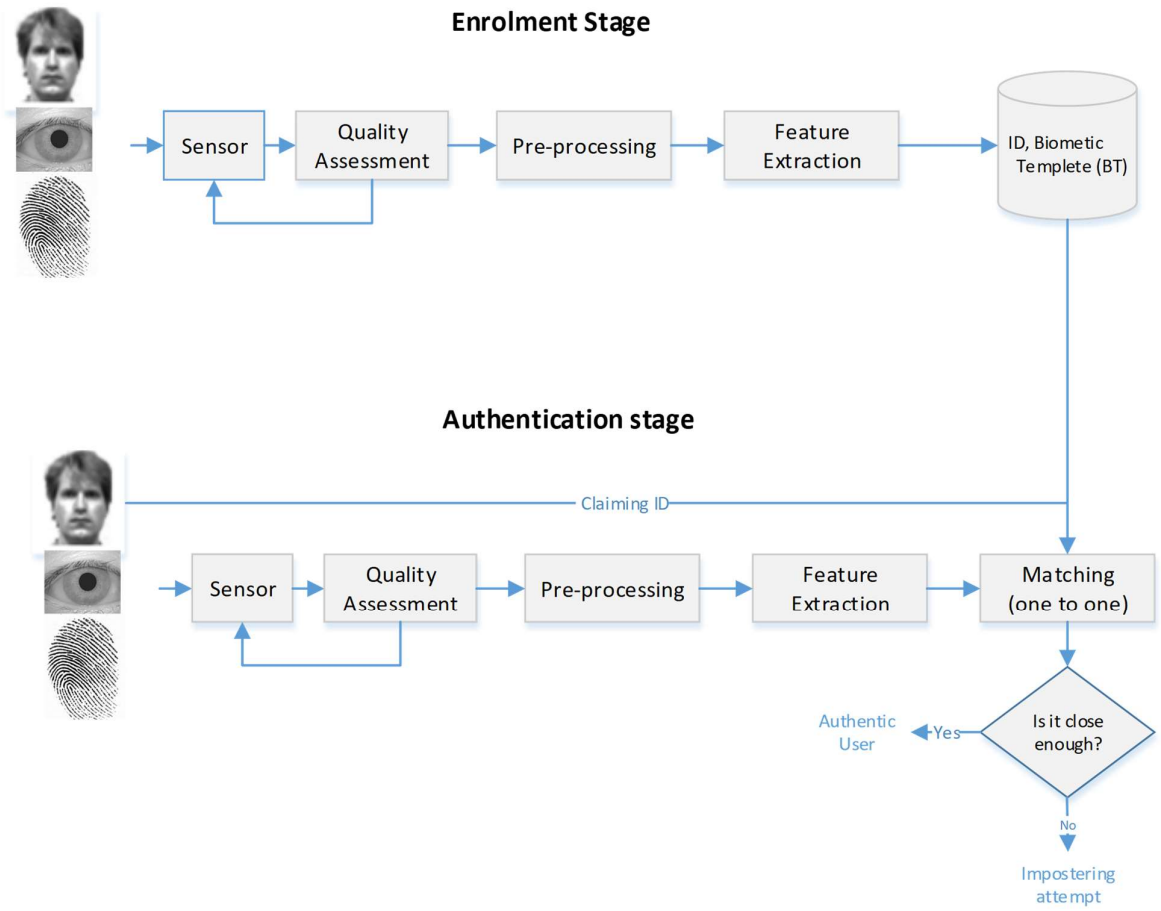

Figure 2. A typical enrolment stage of a biometric system (the face image was used from the Extended Yale Face Database B [13])

At the recognition stage, biometric systems can function in two modes depending on the application context, namely authentication or identification mode.

Biometric-based authentication (also known as verification) is a one-to-one comparison of a freshly captured biometric sample(s), known as query, against an enrolled BT as illustrated in Figure 2. In this mode, a user claims an identity and the biometric system verifies the authenticity of the claimed identity (e.g., the system answers the question: "Are you who you say you are?"). For example, authentication

might be used when a user wants to access his/her bank account or computer. The matching process uses a distance or similarity function to calculate a score indicating the similarity between the stored BT and the fresh feature vector extracted from the query sample. If the matching score is high enough, i.e. close enough to the enrolled template, the biometric system grants access to the user. Otherwise the requested access is rejected. The term "high enough" is determined by the administrator depending on the level of tolerance necessary for the specific application. This allows the system administrator to adjust the rates of false acceptance (i.e. wrongly accepted imposters as genuine users) and false rejection (i.e., wrongly rejected genuine users) to the desired levels. Typically, there is a trade-off between the False Acceptance Rate (FAR) and the False Rejection Rate (FRR), in which the reduction of one rate causes an increase in the other. Most biometric systems are configured to be highly secure by maintaining a very low (e.g. 1 in 10,000) FAR and an acceptable FRR. It is generally less problematic to have a false rejection by asking the genuine user to re-scan their biometric, rather than a false acceptance in which an unauthorized individual will be granted access.

1.4. The limitations of conventional biometric for remote authentication

As we have mentioned previously, many business organisations, government agencies and customers are rapidly shifting many of their services and data onto the cloud which has necessitated the need for secure remote authentication schemes that are immune from fraud and identity theft. Although a biometric-based authentication system is known to be more reliable than traditional authentication schemes, biometric systems can be subject to failure due to the intrinsic factors mentioned earlier or adversary attacks. The security of biometric remote cloud-based authentication in particular can be undermined in a number of ways. For example, a biometric template can be replaced by an imposter's template in the cloud database or it could be stolen and replayed [14]. As a result, the imposter will gain unauthorized access to a place or a system. Moreover, it has been shown that it is possible to create a physical spoof starting from biometric templates [15]. Adler et.al proposed a "hill climbing attack" [16] on a biometric system which in a finite number of iterations, generate a good approximation of the target template. Another method has been also proposed in [17] to reconstruct fingerprint images from standard templates which might fool fingerprint recognition. Furthermore, biometric data on its own is not very secret. Individuals usually unintentionally leave (poor-quality) fingerprints everywhere, and a hidden camera can capture a picture of a face or an iris [18]. In fact, the level of secrecy varies among different biometric modalities (e.g., covertly collecting face images or voice samples is much easier compared to collecting retina and palm vein samples).

Remote biometric authentication in cloud environments is particularly vulnerable to eight possible points of attack highlighted by Ratha et al. [19]. As illustrated in Figure 3, Attack 1 occurs when an attacker presents fake biometric sample at the sensor such as photo of a face, fake fingerprint, copy of a signature, and a recorder voice. Attacks 2 and 4 are replay attacks by resubmitting an old signal by bypassing the sensor or the feature extractor. Attacks 3 and 5 are Trojan horses that produce feature set or matching score

chosen by the attacker. Attack 6 is to target the enrolled templates database stored on the cloud and tamper with the template values. Attack 7 is on the channel between the template database and the matcher where an attacker might tamper with the template before it reaches the matcher. Finally, attack 8 is to override the decision by the attacker.

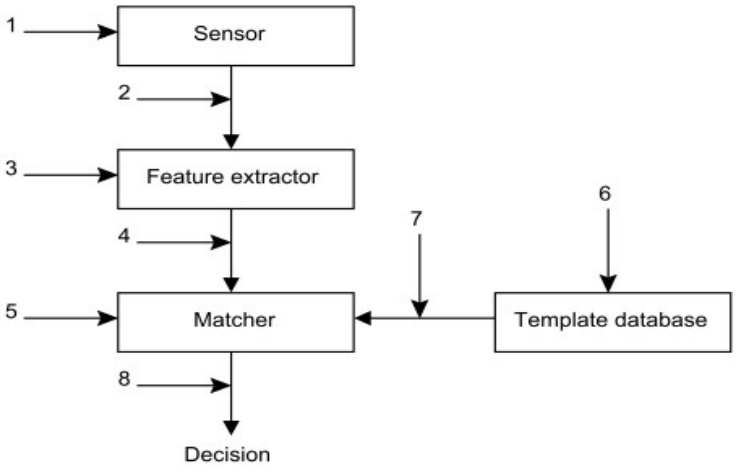

Figure 3. Eight points of attacks in a biometric authentication system (adapted from *[19]*)

Most of the eight points of attacks explained above can be exploited in unattended remote authentication when accessing cloud resource. Therefore, the deployment of such systems for remote authentication is still very limited. Solutions such as cancellable biometric systems and biometric cryptosystems are inherently not immune to replay, man-in-the-middle, and other remote attacks. This necessitate a need for new innovative solutions that help in fraud detection and identity theft prevention in cloud environment as highlighted in section 3.

2. Authentication challenges for the cloud computing environment

With more individuals and organizations opting to use cloud storage service, remote access control to data and cloud resources is becoming a key requirement. This is particularly important as cloud clients do not usually know where their data is physically stored (i.e., the overall security control has shifted form data owners to the hand of service providers [3]). Therefore, the question of how to limit data access to the data owner and authorized user has been one of the biggest challenges in cloud environments. Combination of ID cards and passwords/PINs-based authentication is the most widely used form of remote user authentication [5]. As such authentication factors are not linked with the identity of data owner (see section 1.3), biometric authentication seems the ideal option for access control. However, employing biometric based solutions for cloud authentication is far from a straightforward task due security issues highlighted in section 1.4. Although biometric authentication is perceived to be more reliable than traditional authentication schemes, the open nature of unattended remote authentication makes biometric systems vulnerable to replay and other remote fraudulent attacks. Generally speaking, an effective remote authentication in cloud environments has to be a component of a security package that meets the following requirements [3]:

- The original data should be intractable to restore by cloud providers.
- Offering effective means to distinguish between legitimate users and impostors, and prevent the latter from gaining unauthorized access.
- Preventing any intruder from altering original messages.
- Minimizing response-time to clients' requests, which is a vital requirement for security mechanisms that rely on timestamp synchronization.

2.1. Traditional biometric solutions for cloud authentication

The concept of revocable or cancellable biometric templates generates biometric templates that are not fixed over time. In this case, such templates can be revoked in the same way as lost PINs or passwords [20]. Biometric cryptosystems, on the other hand, aim to generate biometric keys and bio-hashes that are used as a proof of identity instead of biometric templates. In general, the key approaches to protect biometric templates rely on the use of a non-invertible secret transformations on the biometric feature vectors [21].

An interesting implementation of revocable biometrics schemes is the Multi-Factor Biometric Authentication (MFBA) which makes use User-Based Transformations (UBTs) on biometric feature vectors. These MFBA schemes have been proposed to improve security and privacy of biometric systems. Figure 4 shows the key steps of a MFBA system based on the UBT approach at the enrolment and authentication stages. UBTs tend to use transformation keys produced from passwords/PINs or stored on a token.

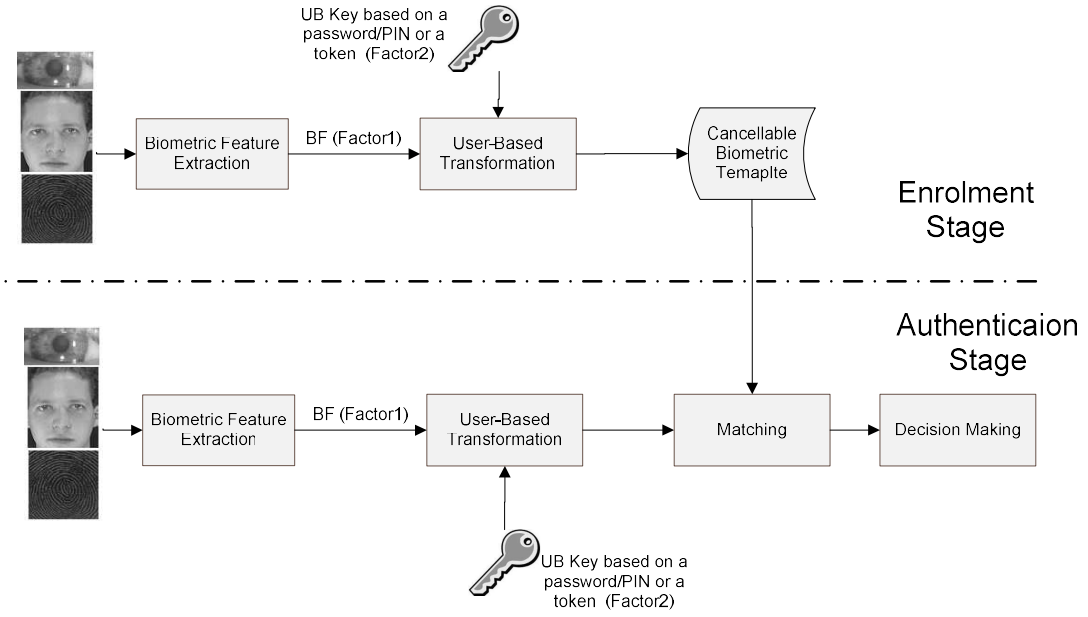

Figure 4. Key steps of a MFBA system based on applying UBTs at the enrolment and authentication stages

It can be argued that revocable biometric systems and biometric cryptosystems enhance the overall security and user's privacy but at the same time they are inherently not robust against replay, man-in-the-middle, and other remote attacks in cloud environments.

Another traditional approach to remote biometric authentication relies on combining MFBA with challenge-response approach between clients and cloud authentication service [22]. As illustrated in Figure 5, such an approach employs a blinding random vector as an essential ingredient of a one-time representation of multi-factor biometric that prevents replay attacks and enhances the security of the MFBA. The production of one-time, fresh revocable biometric template is based on transformations that are generated fresh with contribution from both the client and the cloud authentication server to provide the non-repudiation feature.

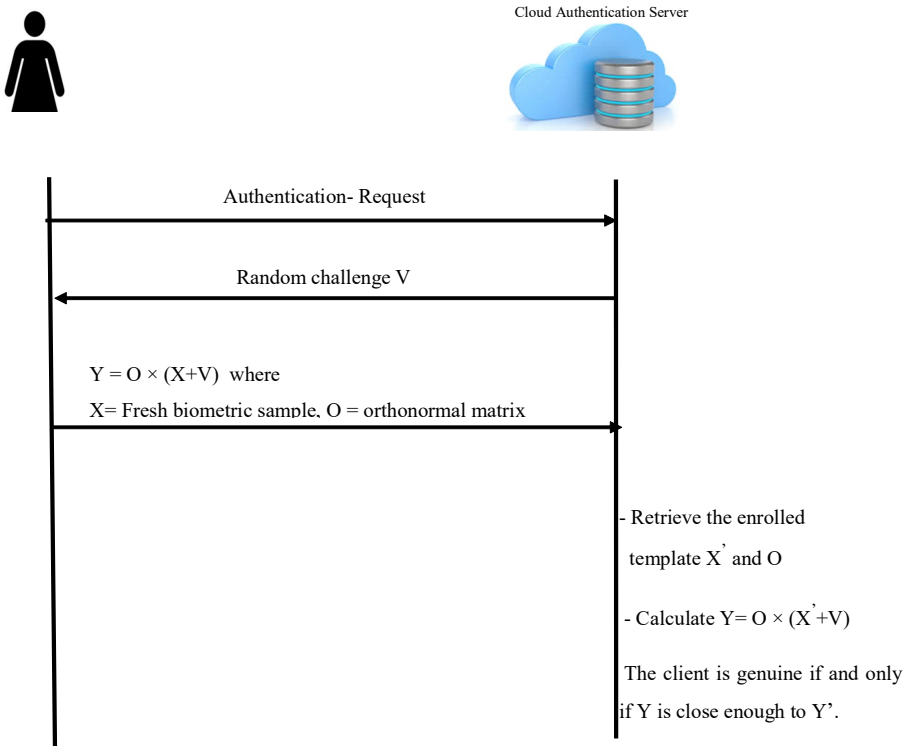

Figure 5. Multi-factor one-time biometric representation based on challenge/response for remote cloud authentication (adapted from [22]])

2.2. Recent solutions to address the security in cloud environment

Next we discuss some recently proposed solutions that provide security in the cloud environment.

- **Attribute-Based Encryption (ABE)**

Data sharing between different parties is a key feature of cloud computing. As discussed earlier, data should be encrypted before being migrated to cloud environments to ensure the security of cloud storage. The main challenge associated with using symmetric/ asymmetric keys encryption is how to securely store and exchange the keys between different parties in an open environment such as cloud computing. The Public Key Infrastructure (PKI) has been providing a practical solution for session key exchange for many web services. The key limitation of the PKI solution is not only the need for a trusted third partly (e.g. certificate authority) but also the missing link between data owner and the encryption keys.

To link encryption keys with users' identities in cloud environments, Attribute Based Encryption (ABE) has been recently proposed as a new formulation of public/private key infrastructure in which the encryption keys are directly derived from a user's identity [23]. In traditional public-key encryption, a message is encrypted for a specific receiver using the receiver's public-key whereas ABE revolutionize

this idea by linking the public key with the identity (e.g. email address) and/or the attribute (e.g. roles) of the receiver. Thus, the key feature of ABE is to enable data owners to share encrypted data with a set of individuals who have a matching set of attributes.

For example, ABE enables the head of a computing department to encrypt a document and share it with members of staff who have attributes {lecturers, admission committee, exam committee}. A threshold to specify the minimum number of required attributes can be used to offer better level of flexibility on who can access the data. In the above example, if the threshold is set to 2, then any staff member with at least two of the three attributes would be able to decrypt the message. The encryption and decryption keys are generated by a trusted third parity based on a set of descriptive attributes as illustrated in Figure 6.

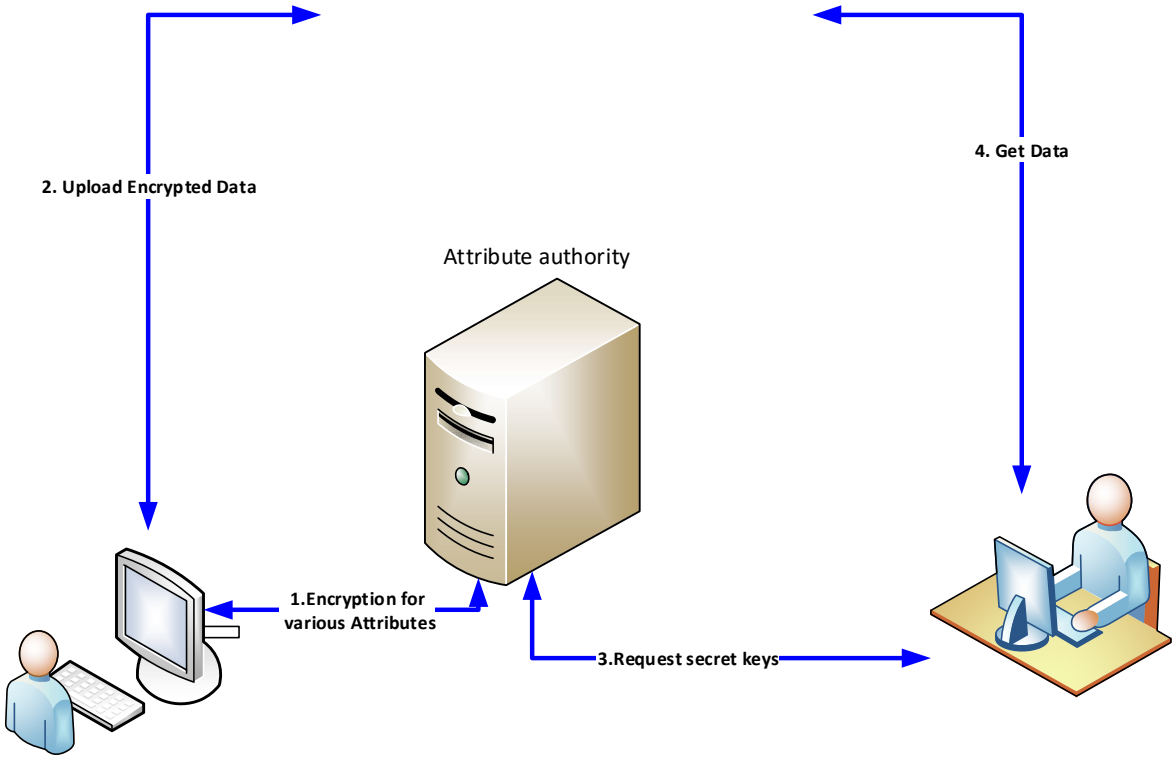

Figure 6. General Attribute-Based Encryption Architecture (adapted from [24])

The following sub-sections further explain the two main extensions of ABE approaches proposed, namely the Key-Policy ABE (KP-ABE) and the Ciphertext-Policy ABE (CP-ABE).

Key-Policy Attribute-Based Encryption (KP-ABE)

The KP-ABE was initially proposed by Goyal et al. [24] to provide a development version of ABE. In KP-APE, data owners generate a master key to encrypt the data in such a way that the corresponding ciphertext is labelled with a set of attributes. The decryption key (private key) given to the user is associated with an access policy (i.e. a tree-like access structure that specifies which ciphertext the key can decrypt). The

leaves of the tree-like access policy are associated with attributes of the users. As a result, a user can decrypt a ciphertext if and only if the attributes associated with the ciphertext satisfy the key access structure.

One application of KP-APE could be the encryption of Audit Log Entries of a big organization. Suppose the entries have the following structure {user name, date and time of action, type of action}, and a forensic analyst is assigned the task of carrying out a particular investigation on the log. If the entries are encrypted using a traditional cryptography, the analyst needs a secret key which will enable him/her to decrypt and access ALL entries. However, in KP-APE, the analyst would be issued a secret key associated with a specific access structure, through which the corresponding decryption key enables a particular kind of encrypted search such as accessing log entries whose attributes satisfied the conditions {"username = John" OR (access date between 01/01/2017 and 01/06/2017)}. The KP-APE also makes it unfeasible for multiple analysts to access unauthorized entries from the audit log even if they collaborate and share their keys [24]. Another simple example of KP-ABE is illustrated in Figure 7.

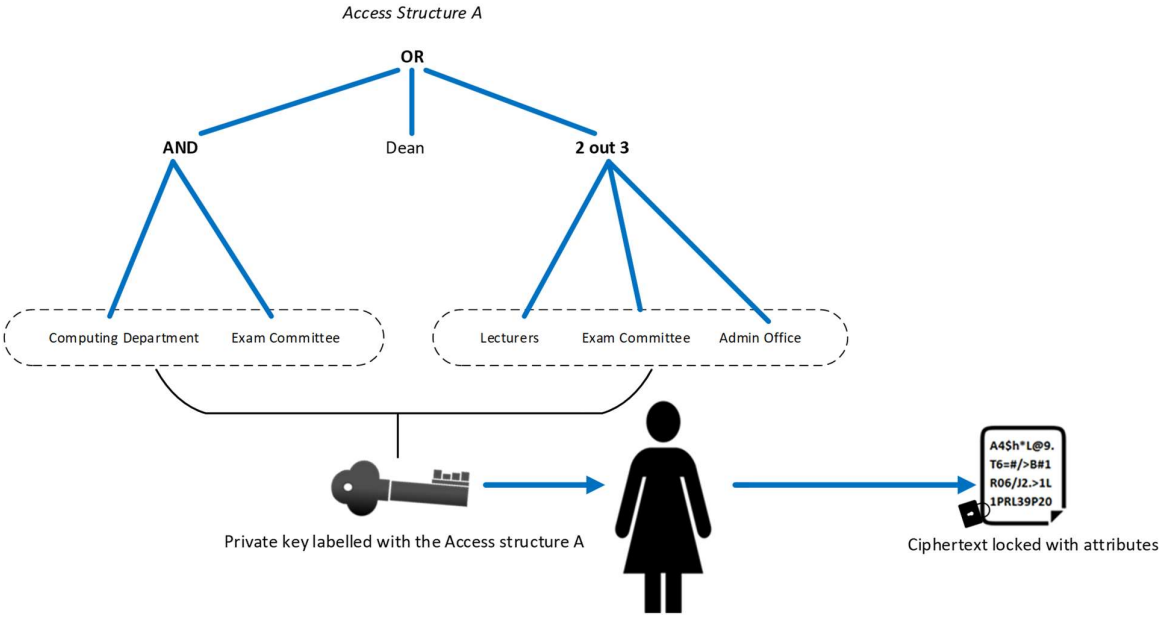

Figure 7. Key-Policy Attribute-Based Encryption model (adapted from [24])

Figure 7 shows a simple access structure which dictates who can retrieve the decryption keys. In this case, Alice would be able to access the decryption key and unlock the ciphertext or part of it if and only if his/her attributes satisfy the corresponding access structure (i.e. she has to be a {Dean OR (a member of Computing Department AND Exam committee) OR (she belongs to two of the three: Lectures, Exam Committee , Admin Office)}).

Ciphertext-Policy Attribute-Based Encryption (KP-ABE)

The main limitation of the KP-ABE is that data owners have no control over who can access the encrypted messages because the access policy which is typically managed by a third party Public Key Generator (PKG) is not attached with the ciphertext (i.e. the access policy controls the access to the decryption keys instead of controlling the access to ciphertext). On the other hand, the Ciphertext-Policy Attribute-Based Encryption (CP-ABE) shifts the focus to the ciphertext by giving data owners the power of locking their encrypted data with different access policies (i.e. for each message they can decide on who can decrypt that particular message as illustrated in Figure 8).

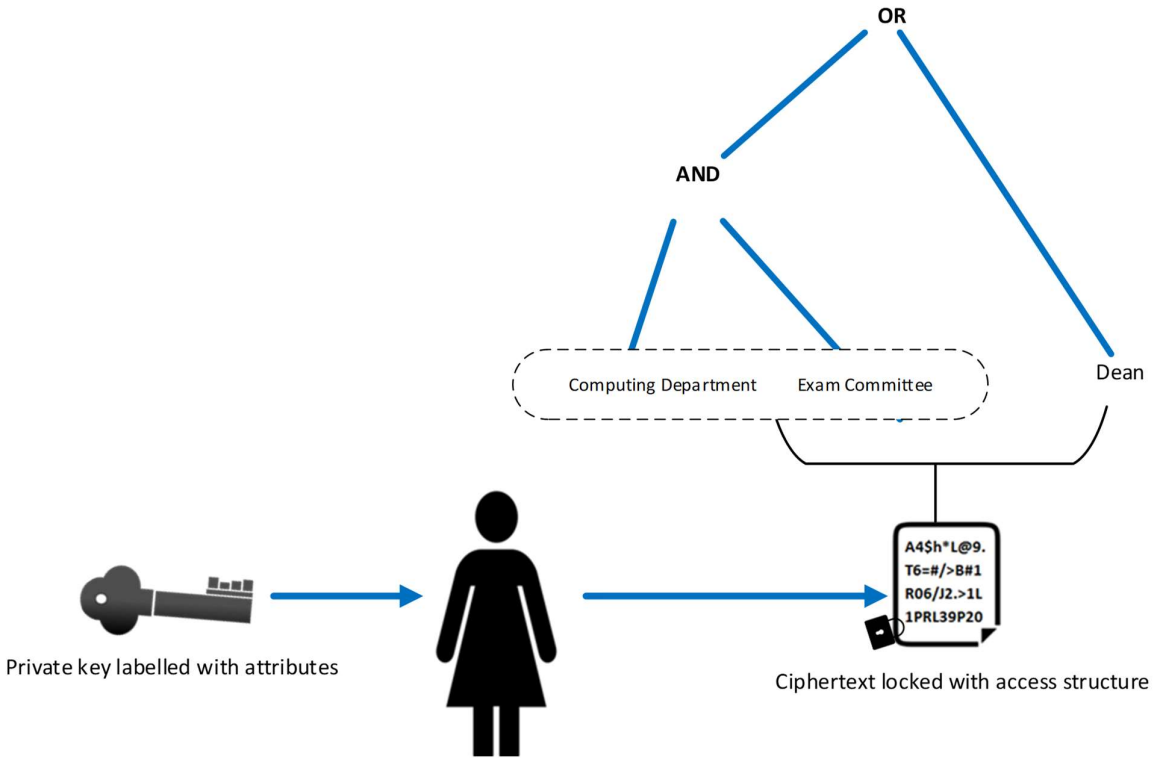

Figure 8. Ciphertext Attribute-Based Encryption Model (adapted from [9])

The example given in Figure 8 demonstrates the flexibility provided to by CP-ABE in locking different messages with different access structures (i.e. data owners are able to choose access policies based on the sensitivity and the security of their encrypted messages). For example, the figure shows that the encrypted message in this scenario can be only decrypted by the Dean OR (a member of both Computing and Examination staff).

Standard Identity-Based Encryption (IBE)

The first proposal of Identity-Based Encryption (IBE) was presented by [25] as an alternative approach to traditional public-key encryption in which the public key of a user is some unique information about the identity of the user such as email address. If Alice wants to send a secure message to Bob, she can use the

text-value of his email address as an encryption key using the public parameters of the system stored on the PKG. The basic concept of IBE can be illustrated as in Figure 9.

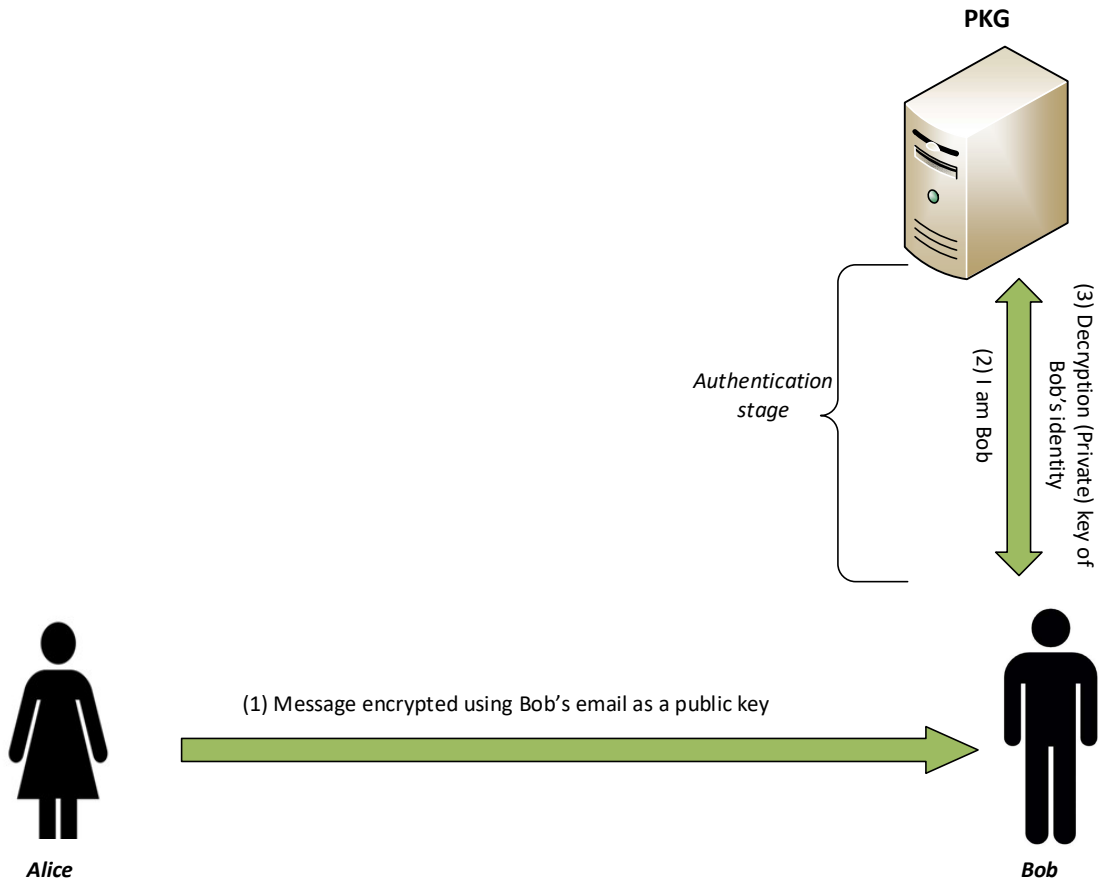

Figure 9. Example of Identity-Based Encryption architecture (adapted from [25])

The figure shows that Alice sends Bob a message encrypted using Bob's Identity such as email, phone number, and so on even if Bob has no public key certificate. The private decryption key is then retrieved from the PKG which holds public parameters as well as the master secret key msk. IBE scheme can offer more power to Alice because she can apply more restrictions on her encrypted message by adding more roles to Bob's ID such as {Bob's email, Access time/data, Role}.

3. Biometric solutions for cloud computing

3.1. Fuzzy Identity-Based Encryption

The initial idea of Fuzzy Identity Based Encryption (F-IBE) was presented in [23] where the identity was modelled as a set of descriptive attributes. The key feature of F-IBE is that the private key of an identity x has an ability to decrypt a ciphertext that has been encrypted with another identity y if and only if the distance between $x\ and\ y$ is less than or equal to a certain threshold value. The F-IBE plays a key role in

utilizing biometric data such as a fingerprints or face images as identity. F-IBE is a promising solution that bridges the gap between the exactness of encryption/decryption keys and the fuzziness of biometric data (i.e. the enrolled biometric samples and the freshly captured ones are never the same). This feature enables a private key of biometric identity to decrypt a message that was encrypted with a public key of a slightly different biometric identity.

It can be argued that the weakness associated with the use of traditional IBE is that the identity such as a "name" or "email" needs to be authenticated first before retrieving the corresponding decryption key in [6] [26] [7]. Therefore, the user might need to provide additional "supplementary documents" to link the name and or the email with her identity. In contrast, the biometric-based F-IBE offers a natural way of authentication by providing biometric data, which is part of user's identity, to retrieve the decryption private key. It has been proved that F-IBE can withstand collusion attacks (i.e. a group of users cannot integrate their keys in a manner that enables them to decrypt messages without individual permission).

In F-IBE, the user's private key is a set of **n** private components or features the are linked within the identity of the user. Shamir's secret sharing [25] is typically employed to distribute the master secret key over the components of the private key by using a polynomial of degree **(d-1)** where **d<=n** is the minimum number of private components that the user needs to present to retrieve the decryption (private) key. Figure 10 shows an example of two identities X and Y where the number of overlapped features is 16 out of 20 features. If the threshold **d** is set to be 16 or less, then the two identities will be deemed to be the same.

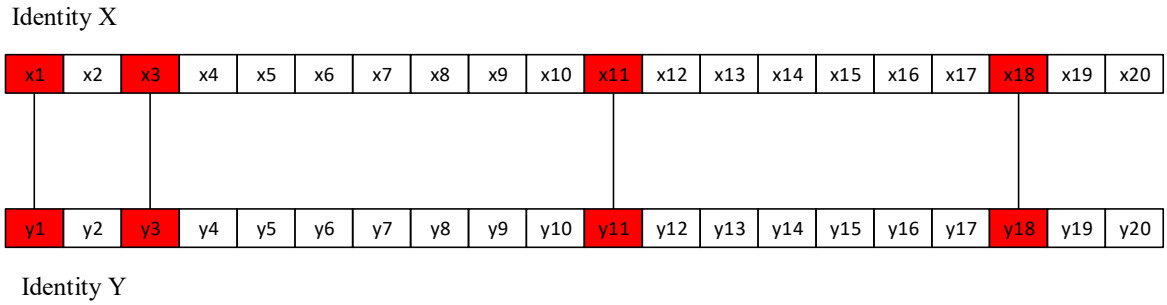

Figure 10. Two identifies X, Y with 16 out of 20 overlap

3.2. Recently proposed biometric solutions

BIOmetric Identity Based Encryption (BIO-IBE) is an interesting application derived from F_IBE, in which the identity x is a feature vector $(f_1, f_2, \dots, f_n)$ of size **n** that is extracted from biometric data using a particular feature extractor technique. On the other hand, the identity **x'** represents a feature vector $(f'_1, f'_2, \dots, f'_n)$ extracted from a fresh biometric query. A polynomial of degree $(\boldsymbol{d}-1)$ is used to distribute and reconstruct the secret value over set of overlapping feature between **x** and **x'** in such away $|\boldsymbol{x} \cap \boldsymbol{x}'| \geq \boldsymbol{d}$. Interpolating the subset of d-features would be enough to reconstruct the secret value and decrypt the ciphertext (encrypted by public key of identity **x**) using the private key components of identity **x'**.

The second application of the F-IBE is ABE, in which a set descriptive attributes are used to encrypt/decrypt a message. For example, the following attributes {Applied computing department, Staff member, age >=40, exam committee member} can be used during encryption. At the decryption stage, anyone has who d-attributes (e.g. 3 out of 4 attribute) should be able to decrypt the message. If d=3, then a person with {Dept =Applied computing, staff member, age= 42} will be able to decrypt the message.

In [27], IBE was developed to build a new scheme of digital signature based on the so-called identity based signature in which the sender uses the recipient's biometric template to encrypt a message. The recipient then authenticate himself/herself to the PKG by presenting a fresh biometric template. The recipient then will be able to decrypt the message if the distance between the two templates is less that a pre-define threshold. In [28], a new protocol for key exchange using biometric identity based encryption was proposed to enable parties to securely exchange cryptographic keys even when an adversary is monitoring the communication channel between the parties. The protocol combines biometrics with IBE in order to provide a secure way to access symmetric keys based on the identity of the users in the unsecure environment.

The message is first encrypted by the data owner using a traditional symmetric key before migrating it to a cloud storage. The symmetric key is then encrypted using public biometrics of the users selected by the data owner to decrypt the message based on Fuzzy Identity-Based Encryption. Only the selected users will be able to decrypt the message by providing a fresh sample of their biometric data. Such a solution eliminates the needs for a key distribution center in traditional cryptography and gives the data owner the power of fine-grained sharing of encrypted data by controlling who can access his/her data.

As illustrated in Figure 10, the key stages of the biometric based IBE framework to exchange keys for sharing encrypted data in the cloud environment [28] are:

- Alice encrypts her data using traditional encryption (symmetric/ asymmetric) techniques such as AES or RSA.

  $$\mathcal{E}_M \longleftarrow Enc(sk, M)$$

  Where sk is the encryption key, *M* is the original message, and $\mathcal{E}_M$ is the encrypted message
- She stores the encrypted data in a cloud environment.
- Now, if Alice wants to allow Bob to decrypt the message, she encrypts the encryption key *sk* using a public key of Bob's unique identity *w′* (i.e., Bob's biometric such as a photo of his face) to produce $\mathcal{E}_{sk}$.

  $$\mathcal{E}_{sk} \longleftarrow Enc(pk_{id}, sk)$$

  Where pk_{id} is the public key of Bob's identity, and $\mathcal{E}_{sk}$ is the encrypted secret key.
- Alice sends the output $\mathcal{E}_{sk}$ to Bob.
- To get the *sk*, Bob needs to provide a fresh biometric sample *w.*
- If and only if the overlap between *w* and *w′* is greater than a threshold value, Bob will retrieve the corresponding private key of his identity and decrypt the ciphertext to get the *sk*.

  $$sk \longleftarrow Dec(sk_{id}, sk)$$

- Bob brings the encrypted data stored in the cloud environment to his local device, and uses *sk* to retrieve the original message/data.

$$M \leftarrow Dec(sk, \mathcal{E}_M)$$

It can be argued that since the face biometric data, for example, is public between parties who know each other, it can be obtained from many resources such as social media resources (e.g., Facebook, Instagram, etc.). Hence, face recognition is an ideal biometric trait for our proposal.

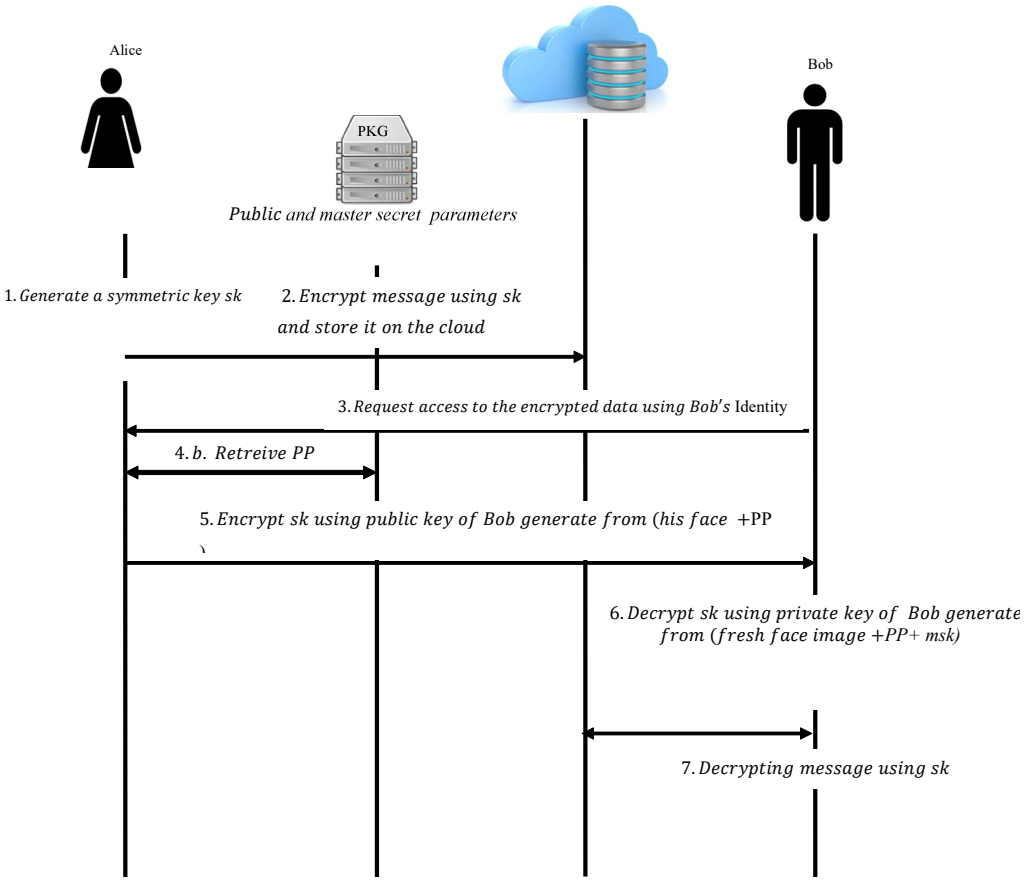

Figure 11. Biometric based IBE framework to exchange keys for sharing encrypted data in cloud environment (adapted from [29])

4. Challenges, Solutions, and Future Research Directions

While the use of biometric data provides a great level of convenience by eliminating the need to remember complex passwords or carrying security tokens, it raises serious questions as to whether, and to what extent, the privacy of the biometric owners can be breached. The fact that we are surrounded by more and more biometric sensors that can capture our biometric traits may eventually limit our privacy in one way or another. The convenience of using biometrics for cloud based authentication could lead to the loss of privacy as a result of being tracked in in many of our daily life activities. Moreover, recent research into biometrics shows that more and more personal information can be revealed from biometric raw data and templates such as gender, age, ethnicity, and even some critical health problems including diabetes, vision

problems, Alzheimer's disease, and so on [30]. Such personal and sometime confidential information might be used to discriminate against individuals when it comes to insurance, jobs, border entry enforcement, and many other issues. Therefore, the challenges in biometric research activities have been expanding to include the maintenance of user's privacy in biometric based systems in addition to the traditional work on enhancing accuracy, scalability, and usability.

In cloud computing, the challenge of incorporating biometrics for robust authentication is not only related to designing a system that is highly accurate and convenient to use but it should also provide an acceptable level of user privacy. The big question here "*do users understand the level of invasion into their privacy as a direct result of improving the security of their data or applications?*" The answer is rather tricky because the acceptable level of privacy invasion is not yet clearly defined in the trade-off between security and privacy. Academic efforts have not stopped over the last few years to come up with a common understanding of the security-privacy trade-off at different levels, which aims to propose best practices and guidelines to national and international policymakers. For example, the SurPRISE (Surveillance, Privacy and Security) project (2012-2015) [31] was an EU funded project that aimed to examine the trade-off between security and individual privacy and addresses the difficult questions of "Does more security justify less privacy?" and "What is the balance between these two?" in which different security-related technologies and measures were evaluated.

4.1. Privacy-Preserving Biometric for Cloud Authentication

Combining challenge/response approach with revocable biometrics generated from UBTs could lead to developing an effective privacy-preserving cloud-based biometric authentication [22].

At the enrolment stage and to address the privacy concerns highlighted earlier, only a revocable biometric features X_{AC} and a hash of a PIN used to generate user-based transformation are stored in the cloud authenticator's database as illustrated in Figure 12. As an example, a client captures an image of his/her face using his/her mobile's camera and key in a 4-digit PIN to produce a UBT. It has been shown that such a transformation improves privacy without compromising on accuracy [22].

During the authentication stage, the client captures a fresh biometric sample and applies the same UBT to produce a revocable feature vector X_C, which is then combined with a one-time random vector V generated by the cloud authentication server. After the output is shuffled using a permutation key generated from the PIN. Due to the variation between the freshly captured sample and the enrolled template, error correcting codes such Reed-Solomon can be used. At the cloud authenticator server, if the error correcting codes was successful, the cloud service produces a key K' that matches the original key if the distance between the biometric template (stored at the cloud authenticator database) and the fresh biometric sample is less than agreed threshold (i.e., they both belong to one client).

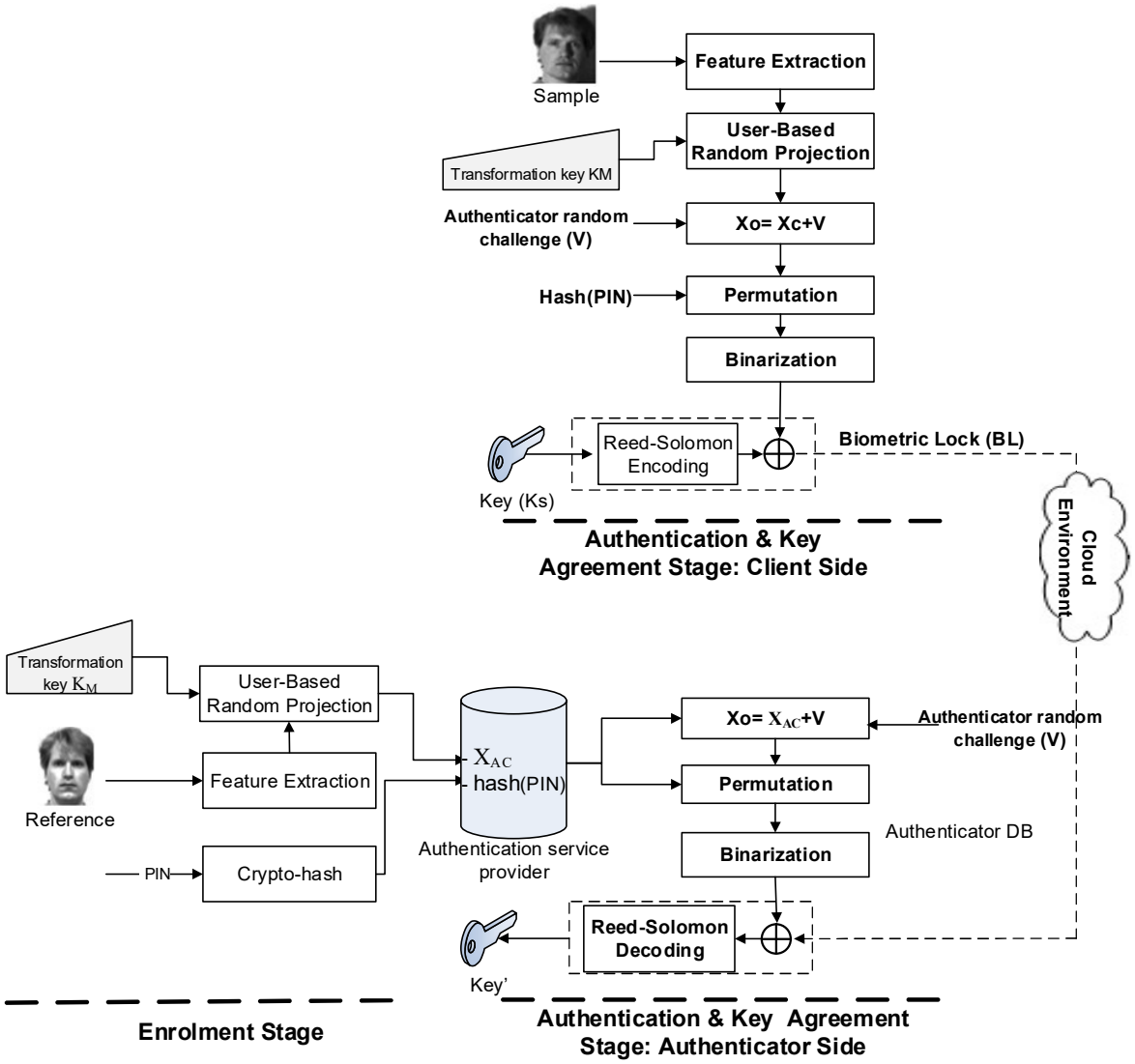

Figure 12. Authentication stage of the privacy-aware authentication scheme for cloud service (adapted from [22])

4.2. Conclusion and future research direction

While biometric-based cloud authentication can provide a convenient way to manage access control to cloud resources, biometric data can be misused to track individuals and leak confidential information related to health, gender, ethnicity, and so on. It can be argued that privacy of cloud-based biometric authentication cannot be solely addressed by technical means. Therefore, there is an urgent need for new legislation to enforce privacy-aware measures on cloud service providers related to biometric collection, data processing, and template storage. Although some types of regulation related to users' privacy and data protection do exist in many countries, many of these regulations related to managing the privacy and

security of biometric data are either not there yet or insufficient. In the end, technical solutions have to complement legal frameworks to enforce certain measures on cloud authentication services, which would eventually lead to wider public acceptance of biometric based solutions.

References

[1] W. Jansen, T. Grance and others, "Guidelines on security and privacy in public cloud computing," *NIST special publication,* vol. 800, no. 144, pp. 10-11, 2011.

[2] S. Dustdar, "IEEE Computer," *Cloud Computing,* vol. 49, no. 2, pp. 12-13, 2016.

[3] J. W. a. J. F. R. Rittinghouse, Cloud computing: implementation, management, and security, CRC press, 2016.

[4] H. Takabi, J. B. Joshi and G.-J. Ahn, "Security and privacy challenges in cloud computing environments," *IEEE Security \& Privacy,* no. 6, pp. 24-31, 2010.

[5] Z. Xiao and Y. Xiao, "Security and privacy in cloud computing," *Communications Surveys \& Tutorials, IEEE,* vol. 15, no. 2, pp. 843-859, 2013.

[6] D. Boneh and M. Franklin, *SIAM Journal on Computing,* vol. 32, no. 3, pp. 586-615, 2003.

[7] D. Boneh and X. Boyen, in *International Conference on the Theory and Applications of Cryptographic Techniques*, 2004.

[8] X. Boyen and B. Waters, "Anonymous hierarchical identity-based encryption (without random oracles)," in *Advances in Cryptology-CRYPTO 2006*, Springer, 2006, pp. 290-307.

[9] J. Bethencourt, A. Sahai and B. Waters, "Ciphertext-policy attribute-based encryption," in *Security and Privacy, 2007. SP'07. IEEE Symposium on*, 2007.

[10] M. Chase, "Multi-authority attribute based encryption," in *Theory of cryptography*, Springer, 2007, pp. 515-534.

[11] N. Cao, C. Wang, M. Li, K. Ren and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," *Parallel and Distributed Systems, IEEE Transactions on,* vol. 25, no. 1, pp. 222-233, 2014.

[12] A. A. R. a. K. N. A. Jain, Introduction to biometrics, Springer, 2011.

Al-Assam, H., Hassan, W. and Zeadally, S., 2019. Automated Biometric Authentication with Cloud Computing. In Biometric-Based Physical and Cybersecurity Systems (pp. 455-475). Springer, Cham.

[13] Georghiades, AS., P. N. Belhumeur, and D. J. Kriegman, "From Few to Many: Generative Models for Recognition under Variable Pose and Illumination," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 23, no. 6, p. 643–660, 2001.

[14] Jain, AK. , Nandakumar, K. and Nagar A., "Biometric Template Security," in *EURASIP Journal on Advances in Signal Processing*, 2008.

[15] Nandakumar, K., "Multibiometric Systems: Fusion Strategies and Template Security," PhD thesis, Michigan State University, 2008.

[16] Adler, A., "Vulnerabilities in biometric encryption systems," in *Proc. of the 5th Int Conference on Audio and Video-Based Biometric Person Authentication*, 2005.

[17] Cappelli, R. , Lumini, A., Maio, D., and Maltoni, D., "Fingerprint Image Reconstruction from Standard Templates," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007.

[18] Hao, F. and Anderson, R. and Daugman, J, "Combining cryptography with biometrics effectively," *IEEE Transactions on Computers,* pp. 1081--1088, 2006.

[19] Ratha, N.K. Connell, J.H. and Bolle R.M., "An analysis of minutiae matching strength," *Proc. of Third International Conference on Audio- and Video-Based Biometric Person Authentication,* pp. 223-228, 2001.

[20] H. S. S. J. Hisham Al-Assam, "A lightweight approach for biometric template protection," pp. 73510P-73510P-12, 2009.

[21] S. J. Hisham Al-Assam, "Security evaluation of biometric keys," *computers & security,* vol. 31, no. 2, pp. 151-163, 2012.

[22] S. J. Hisham Al-Assam, "Robust Biometric Based Key Agreement and Remote Mutual Authentication," in *The 11th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, Liverpool, UK, 2012.

[23] A. Sahai and B. Waters, "Fuzzy identity-based encryption," in *Advances in Cryptology--EUROCRYPT 2005*, Springer, 2005, pp. 457-473.

[24] F. Li, "Context-Aware Attribute-Based Techniques for Data Security and Access Control in Mobile Cloud Environment," 2015.

[25] A. Shamir, "Identity-based cryptosystems and signature schemes," in *Advances in cryptology*, 1984.

[26] R. Canetti, S. Halevi and J. Katz, "A forward-secure public-key encryption scheme," in *Advances in Cryptology—Eurocrypt 2003*, Springer, 2003, pp. 255-271.

Al-Assam, H., Hassan, W. and Zeadally, S., 2019. Automated Biometric Authentication with Cloud Computing. In Biometric-Based Physical and Cybersecurity Systems (pp. 455-475). Springer, Cham.

[27] B. Waters, "Efficient identity-based encryption without random oracles," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, 2005.

[28] H. A.-A. Waleed Hassan, "Key Exchange Using Biometric Identy Based Encryption for sharing encrypted data in cloud environment," in *SPIE Defense, Security, and Sensing*, 2017.

[29] W. K. Hassan and H. Al-Assam, "Key exchange using biometric identity based encryption for sharing encrypted data in cloud environment," in *Proc. SPIE 10221, Mobile Multimedia/Image Processing, Security, and Applications*, 2017.

[30] K. E. Forbes, M. F. Shanks and A. Venneri, "The evolution of dysgraphia in Alzheimer's disease," *Brain research bulletin,* vol. 63, no. 1, pp. 19-24, 2004.

[31] [Online]. Available: http://www.surprise-project.eu/.

[32] V. Goyal, O. Pandey, A. Sahai and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," in *Proceedings of the 13th ACM conference on Computer and communications security*, 2006.

[33] R. Ostrovsky, A. Sahai and B. Waters, "Attribute-based encryption with non-monotonic access structures," in *Proceedings of the 14th ACM conference on Computer and communications security*, 2007.

[34] D. Boneh, G. Di Crescenzo, R. Ostrovsky and G. Persiano, "Public key encryption with keyword search," in *International Conference on the Theory and Applications of Cryptographic Techniques*, 2004.

[35] S. Bradshaw, C. Millard and I. Walden, "Contracts for clouds: comparison and analysis of the Terms and Conditions of cloud computing services," *International Journal of Law and Information Technology,* vol. 19, no. 3, pp. 187-223, 2011.

[36] T. Kuseler, H. Al-Assam, S. Jassim and I. A. Lami, "Privacy preserving, real-time and location secured biometrics for mCommerce authentication," in *Mobile Multimedia/Image Processing, Security, and Applications 2011*, SPIE, Bellingham, WA, 2011.

[37] P. Yang, Z. Cao and X. Dong, "Fuzzy Identity Based Signature.," *IACR Cryptology EPrint Archive,* vol. 2008, p. 2, 2008.

[38] N. D. Sarier, "A new biometric identity based encryption scheme secure against DoS attacks," *Security and Communication Networks,* vol. 4, no. 1, pp. 23-32, 2011.

[39] R. Sakai and M. Kasahara, "ID based Cryptosystems with Pairing on Elliptic Curve.," *IACR Cryptology ePrint Archive,* vol. 2003, p. 54, 2003.

[40] T. G. W. Jansen, "Guidelines on security and privacy in public cloud computing," *NIST special publication,* vol. 800, no. 144, pp. 10-11, 2011.

Al-Assam, H., Hassan, W. and Zeadally, S., 2019. Automated Biometric Authentication with Cloud Computing. In Biometric-Based Physical and Cybersecurity Systems (pp. 455-475). Springer, Cham.