# BIOVAULT: BIOMETRICALLY BASED ENCRYPTION

Mr.B.L. Tait[1], Prof S.H. von Solms[2]

[1] University of Johannesburg, Kingsway Avenue, Auckland Park, Gauteng, South Africa, Btait@uj.ac.za
[2] University of Johannesburg, Kingsway Avenue, Auckland Park, Gauteng, South Africa, basie@uj.ac.za

**Abstract.** Biometric based characteristic authentication is an asymmetric [1] authentication technology. This means that the reference biometric data generated during the enrolment process and stored in the biometric database, will never match any freshly offered biometric data exactly (100%). This is commonly accepted due to the nature of the biometric algorithm [2] central to the biometric environment.

A password or pin on the other hand, is a symmetric authentication mechanism. This means that an exact match is expected, and if the offered password deviates ever so slightly from the password stored in the password database file, authenticity is rejected.

Encryption technologies rely on symmetric authentication to function, as the password or pin is often used as the seed for a random number that will assist in the generation of the cipher. If the password used to encrypt the cipher is not 100% the same as the password supplied to decrypt, the cipher will not unlock.

The asymmetric nature of biometrics traditionally renders biometric data unfit to be used as the secret key for an encryption algorithm.

This paper introduces a system that allows biometric data to be used as the secret key in an encryption algorithm. This method relies on the BioVault infrastructure. For this reason BioVault will briefly be discussed, followed by a discussion of biometrically based encryption.

**Keywords:** Encryption, Biometrics, BioVault, security, secure transaction, data protection, key management, privacy-enhancing technology, data security.

# 1 Introduction.

To date, it was not possible to use a biometric data directly as the secret key for an encryption algorithm or for a MAC algorithm. The reason for this resides in the fact that a biometric authentication process is always asymmetric. In order for an encryption algorithm to function, the secret key provided to encrypt a message must be exactly the same (symmetrical) as the secret key used to decrypt the message.
If a secret key is used to generate a MAC, this exact same secret key must be provided to test the MAC.
The possibility that a person would repeatedly be able to provide biometric data that would be 100% the same as earlier provided biometric data is highly unlikely. This makes biometric data useless as the secret keys for hashing or encryption.

Digital signatures use encryption and hashing as its underlying, primary technology.

The paper is based on the BioVault protocol. Because of the length restriction on this paper, the BioVault protocol cannot be discussed in detail. However a short discussion of the protocol will be given, followed by a detailed discussion of how the protocol can be used to create biometrically based digital signatures.

For a detailed discussion of the BioVault protocol see [6], [7].

In the sections to follow it will be demonstrated how the BioVault infrastructure allows biometric data to be used for encryption.

# 2 Encryption using a secret key or biometric data.

## 2.1 Secret key encryption.

If a user wishes to send a message to another user over an unsecured network, the message must be encrypted in one or other way. C = E*k(M) [3] where:
C = Cipher message
M = Original message
k = secret key
E = Encryption algorithm

The typical encryption process using a secret key is illustrated in figure 1

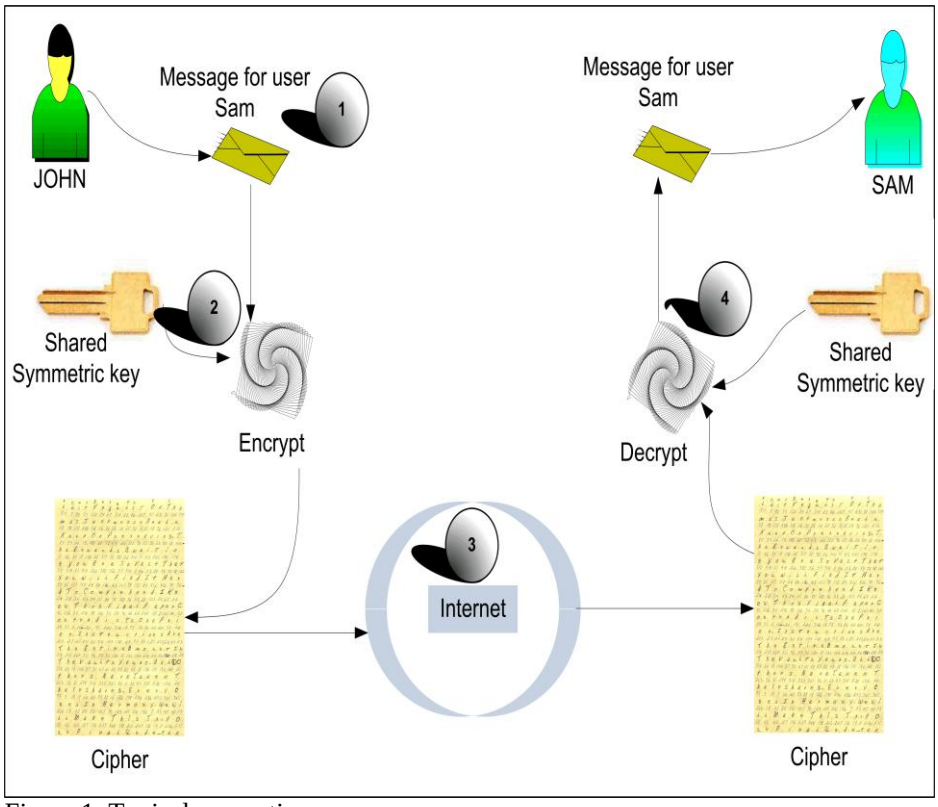


Figure 1: Typical encryption process.

As illustrated in figure 1, John wishes to send a secret message to Sam. In order to secure the message during the transmission, John encrypts the message using an encryption algorithm. In order for the encryption algorithm to yield cipher text that is absolutely random, a secret key must be provided. This secret key is shared between Sam and John as illustrated in figure 1. The secret key provided by John to encrypt the message is exactly the same as the secret key that Sam will provide to decrypt the message.

*Step 1*
John generates the message that he wishes to send to Sam.

*Step 2*
John provides a secret key to the encryption algorithm, and the encryption algorithm uses this secret key to generate the cipher text.

*Step 3*
The message in cipher text is sent over the internet to Sam. If a hacker should intercept this message, the hacker must be in possession of the secret key shared between Sam and John, in order to decrypt the message.

*Step 4*
Sam receives the message sent by John and uses the same encryption algorithm and the secret key that is shared between the two of them. If the secret key that Sam supplied to the encryption algorithm is exactly the same as the secret key used by John, Sam will retrieve the original, unencrypted message that John created.

From the above mentioned example it becomes clear that biometric data can not be used for secure encrypted communication between two people.
If John used his biometric characteristic as the secret key for encrypting a message destined for Sam, Sam would not be able to provide the same biometric characteristic to decrypt the message (as this was John's biometric characteristic that Sam does not possess).

In this paper it is illustrated in what way the BioVault infrastructure is a solution in allowing John to send an encrypted message to Sam, by using his biometric characteristic. This method relies on the fact that both John and Sam are part of the BioVault infrastructure – very much as EBay [4] relies on the fact that buyers and sellers are both part of the PayPal [5] environment. The BioVault infrastructure is a new development, and subsequently not commonly known. For this reason the following section will give a brief outline of the BioVault infrastructure, followed by an explanation of biometrically based encryption. For a detailed discussion of the BioVault infrastructure see [6], [7].

## 3 Brief introduction to BioVault version 3.0

BioVault does not rely on any specific biometric technology to function, however certain technologies are inherently stronger technologies and would obviously be preferred by industry.

During the development of the BioVault protocol the following important goals were set:

1. Safe transport of biometric data over an un-safe network like the internet.
2. Detection of replay attempts of biometric data in electronic format.
3. Protection against manufactured biometric characteristics from latent prints.
4. Enabling a user to use biometric data to encrypt a document
5. Enabling a user to use a biometric data to digitally sign a document.

Enabling a user to use biometric data to digitally sign a document (5), will not be discussed in this paper.

### 3.1 Symmetry and Asymmetry.

One of the fundamental concepts of the BioVault protocol relies on the fact that the biometric authentication process is inherently asymmetric. This makes virtually every presented biometric characteristic unique. This feature that is inherent to biometric technology can be used to detect any form of electronic replay of earlier presented biometric data. A 100% match between the reference biometric data stored in the biometric store, and the biometric data presented by the user, is unlikely. Furthermore it is possible to record biometric data, and check if any biometric data was ever received before.

Password and token based authentication mechanisms, on the other hand, are symmetric. Whenever symmetric mechanisms are to be used, the fact remains that a symmetric match must be absolutely symmetric, thus a 100% correlation is expected between the stored password in the password database, and the presented password.

### 3.2 BioVault components.

The following components are part of the BioVault infrastructure:

#### 3.2.1 The Bio-Archives (BA).

Two Bio-archives (BA) are created; one bio-archive on the authentication server known as the Server bio-Archive (SBA) and one Client side bio- Archive known as the CBA. The SBA will store all biometric data used by the user that was successfully authenticated by the biometric matching algorithm. The SBA will assist in the identification of possible replay attacks. For this reason access to the biometric data stored in the SBA must be very fast. To ensure that specific biometric data inside the SBA can be found very fast, the SBA will be sorted. Considering that SBA is sorted, a binary search algorithm can be used to find biometric data in the SBA efficiently.

The CBA will assist in biometric data protection during transmission.

Initially the CBA will consist of a limited number of previously used biometric data of the specific user (to be discussed in more detail later). The larger this bio-archive the stronger the system will be.

The biometric data inside this CBA are totally random and provided to the user by the authentication server. The authentication server will populate the CBA from time to time with different previously offered biometric data of the given user.

Whenever a secure connection is established between the user and the authentication server, the server can update the CBA. However it is recommended that the CBA is updated under strictly controlled environments. This means that CBA can be updated by the authentication server, whenever a user visits a bank or ATM machine, as an example.

*CBA storage.*

The Bio-Archive that the user will use, will store previously offered biometric data. The following are possible options that can be used to store the CBA.

1. A USB flash memory – These tiny appliances like the Micro SD memory, presently offer surprisingly large storage space with storage sizes reaching 64Gb [114], furthermore, no additional equipment will be needed to integrate this technology into the environment.
2. A Smart card –These devices however need additional equipment and storage capacity on smartcards is limited.
3. A subcutaneous microchip – This technology ensures that a person cannot forget or misplace his CBA, but workable and acceptable solutions are still in development. Storage capacity is limited and technology is controversial. [8], [9].

### 3.2.2 The Bio- Parcel used during the authentication process.

The Bio- parcel will always include freshly offered biometric data and old biometric data that is obtained from the CBA as requested by the Authentication server. The contents of the bio-parcel will be joined using a XOR operator. This is illustrated in figure 2. The aim of the XOR operator is to secure the bio-parcel while transmitted over a public network, without using encryption systems. Encryption systems using for example shared symmetric keys, introduces a lot of system overhead.

For the example as illustrated in figure 2 the CBA would include 50 randomly picked biometric tokens from the SBA of this specific user. The SBA on the server will still include each and every biometric data ever used by the user in his lifetime. As will soon be discussed, these randomly selected biometric data of the user, will serve as a special key, and can be compared to the working of a one time pad

## 3.3 BioVault Mechanism

*Step 1* (As in figure 2)
When a user needs to be authenticated the user attaches the appliance containing the CBA with the previously offered biometric data to the terminal (for example the user's computer or ATM machine), where he intends to do the transaction.

*Step 2*
The user provides a fresh biometric characteristic as shown, directly to the biometric scanner. The scanner will digitize the biometric characteristic and forward the biometric data to the driver software of the biometric device.

*Step 3*
During the previous encounter with the authentication server, the server sent a challenge to the. This challenge demanded specific biometric data from the CBA that

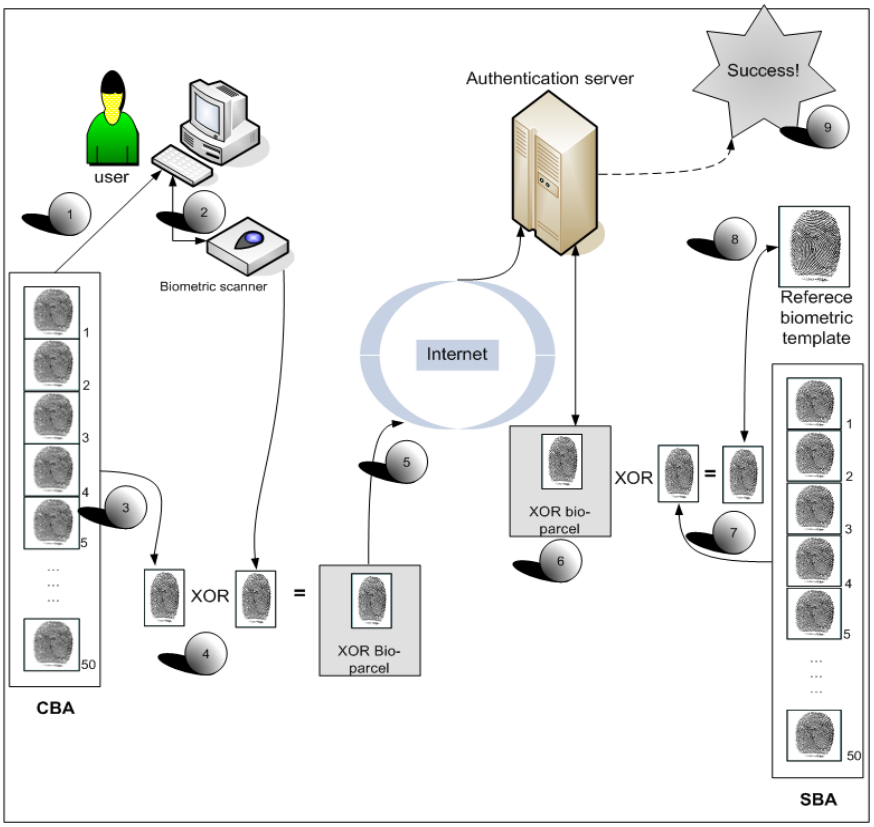


Figure 2: BioVault version 3.0.

had to be included at the time of the next contact with the authentication server. In figure 2, the server requested the 4th biometric data in the CBA. The system will thus automatically obtain the 4th biometric data from the user's CBA.

*Step 4*
The BioVault client side software will take the electronic representation of the freshly offered biometric data and XOR it with the electronic representation of the 4th biometric data obtained in step 3 from the CBA. For example:

Electronic representation of fresh biometric data from scanner: 10101110111011010
Electronic representation of challenged (4th) data from CBA: 10110101111011110
New bio-archive after XOR process: 00011011000000100

This result in a smaller bio-parcel than proposed in BioVault version 2.0, as only the result of the XOR process will be submitted to the authentication server as the XOR bio-parcel.

*Step 5*
The XOR bio-parcel is submitted via the internet or any networked environment to the authentication server.

*Step 6*
The server receives the XOR bio-parcel as shown in step 6, and prepares to run the XOR operator on the bio-parcel.

*Step 7*
The server requested previously that the client XOR the fresh biometric data with the fourth biometric data in the CBA. The server obtains the biometric data in the SBA that corresponds with the expected biometric data received from the user in the XOR bio-parcel.
The server must then XOR the received XOR bio-archive with the 4th biometric data from the SBA, corresponding with the 4th biometric data in the CBA, in order to get the fresh biometric data of the user. For example:

| | |
|---|---|
| XOR bio-archive received from user: | 00011011000000100 |
| Expected 4th biometric data from SBA: | 10110101111011110 |
| Result of XOR process = the fresh biometric data: | 10101110111011010 |

*Step 8*
The fresh biometric data extracted from the XOR bio-archive during step 7, is now asymmetrically matched to the reference biometric template found in the database. The authentication server compares the freshly offered biometric data with the reference biometric template. If the offered biometric data falls within the tolerances defined in the matching algorithm, the system declares the biometric data as authentic and adds this biometric data to the SBA, after checking the SBA for an exact match.

*Step 9*
As the bio-parcel passed all the requirements, authentication is pronounced successful. The server will proceed to the generation of a new challenge destined for the user.

# 4 Biometric Encryption using BioVault.

The whole encryption method using the BioVault infrastructure is a 4-phased process.

## 4.1 Biometric encryption overview.

In phase 1, John identifies himself to the authentication server, and indicates that he wants to send an encrypted message to Sam. In order to send an encrypted message to Sam, John requests a “biometric key” of Sam from the server.
In phase 2, the authentication server retrieves a biometric key from Sam’s STA also found in Sam’s CTA, and sends it to John
In phase 3, John uses this biometric key of Sam, as an encryption key to create the encrypted message, and sends this encrypted message to Sam over the network.
In phase 4, Sam receives the message sent by John, and decrypts the message by testing all the biometric keys in her CTA, against the received cipher text. In essence, Sam will the ‘brute force’ the decryption of the cipher.

## 4.2 Biometric encryption discussion.

Figure 3 illustrates the first phase that John would follow in order to send an encrypted message to Sam.

### 4.2.1 Request of biometric data

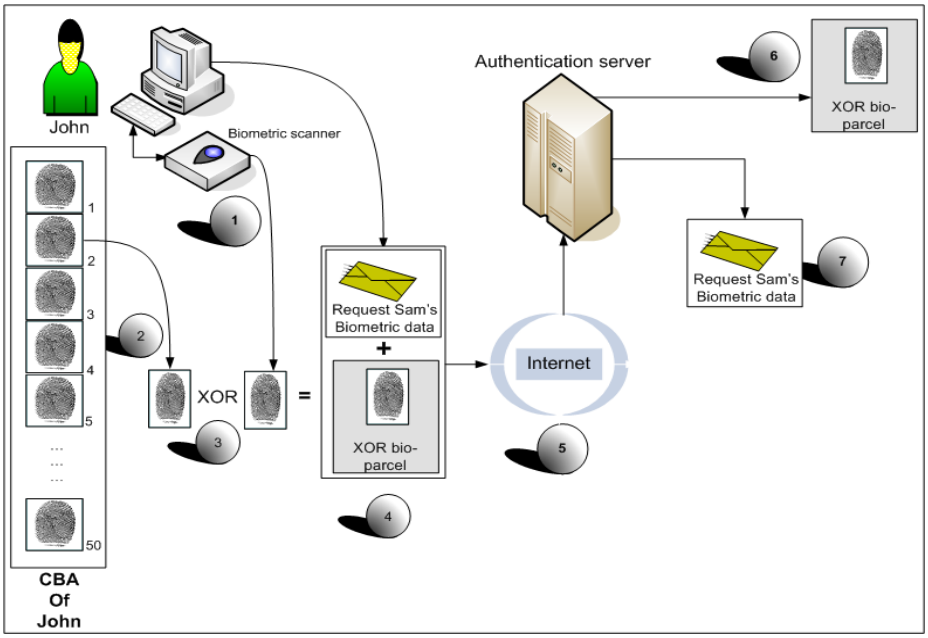


Figure 3: Request biometric data.

At this stage John sent a request to the server, stating that he wished to communicate with Sam. The server authenticated John, based on the fact that the fresh biometric data supplied by John was accepted and the expected biometric data from John's CBA was correctly supplied.

Subsequently the server ensured that Sam is a user on the BioVault system, allowing the second phase to commence. Phase two is illustrated in figure 4.

### 4.2.2 Phase 2: Submission of biometric data of Sam to John

During the second phase the server sends stored biometric data from the SBA of Sam, back to John. The server is aware that this biometric data exists inside Sam's CBA. The steps below explain this process:

*Step 1*
The server obtains biometric data, in this particular illustration the second biometric data, from the SBA of the user Sam. This biometric data is also present in the CBA of user Sam.

The server marks this biometric data as "used for encryption" to prevent this particular biometric data ever again rendered for encryption or authentication. This guarantees that Sam and John are the only people in possession of this biometric data.

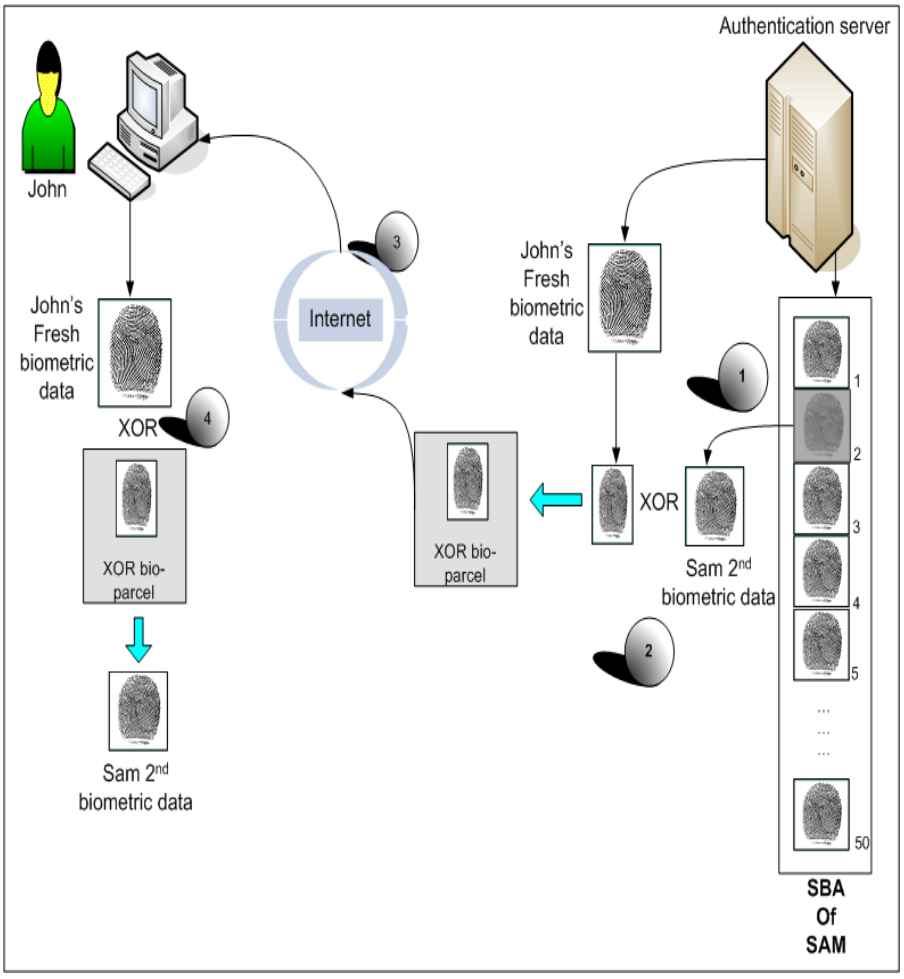

Figure 4 Submission of Sam's biometric data to John:

*Step 2*
The server will XOR the biometric data from Sam's SBA, in this case the 2nd one, with the fresh biometric data received in phase 1 from John, creating a new XOR bio-parcel.

*Step 3*
The XOR bio-parcel is then transmitted via the network, back to John. If this parcel is sniffed during transmission, the hacker will not have much use for the received bio-parcel.

*Step 4*
John receives the XOR bio-parcel. John uses the fresh biometric data he supplied during the first phase, and XOR this fresh biometric data with the bio- parcel received. This step yields the biometric data sent by the authentication server to John – i.e. biometric data number 2 in Sam's CBA.

Once John is in possession of this biometric data of Sam, John can proceed to the third phase, of sending an encrypted message to Sam.

### 4.2.3 Phase 3: Encrypted communication between John and Sam.

At this stage John is in possession of a symmetric copy of the second biometric data in the CBA of Sam. He can proceed to encrypt a message for Sam using the biometric data made available by the server of biometric data found in Sam's CBA, as illustrated in figure 5.

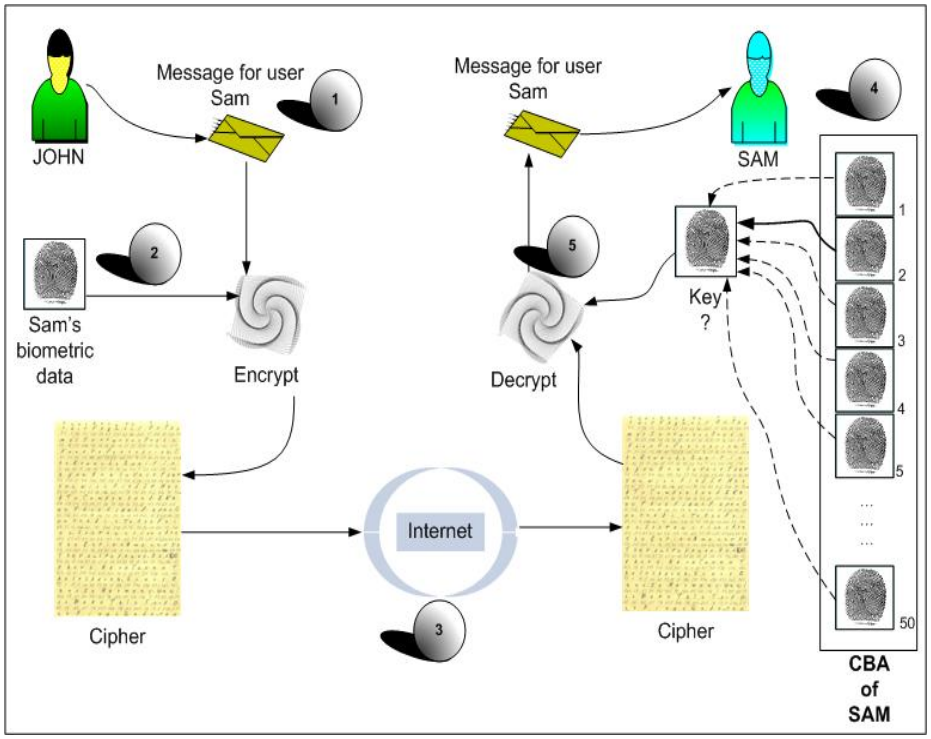

Figure 5: Encrypted communication between John and Sam.

It is illustrated in figure 5 the following steps indicates how John will send an encrypted message to Sam.

*Step 1*
John generates the message that he intends to send to Sam.

*Step 2*
John provides the received biometric data of Sam to the encryption algorithm, and the encryption algorithm uses this biometric data as a secret key to generate the cipher text.

*Step 3*
The message in cipher text is sent via the internet to Sam. If a hacker should intercept this message, the hacker must be in possession of the correct biometric data of Sam, in order to decrypt the message. Considering the working of BioVault version 3.0, this is highly unlikely.

In the final phase Sam will need to decrypt this message sent by John to her, using the biometric data inside her CBA. This process is illustrated in step 4 and step 5 of figure 5.

*Step 4*
Sam receives the message sent by John and accesses her own CBA. The client software on Sam's machine uses all the biometric data in her CBA to brute force the cipher. As there are only a limited number of biometric data in the CBA, this process will unlock the cipher rapidly.

*Step 5*
As the biometric data Sam used to decrypt the message is the same as the biometric data used by John, Sam will retrieve the original, unencrypted message created from the cipher created by John.

## 5 Conclusion.

This paper demonstrated that the BioVault infrastructure makes it absolutely possible to encrypt a message using biometric data.

Biometric data relates directly to the users. If a user used a person's biometric characteristic to encrypt a message (similar to using a person's public key in the PKI system) only the receiving party with the correct biometric data will be able to decrypt the message- however unlike the PKI system, biometric data is directly related to the user. If tokens and passwords are used, only the token or password are authenticated, the user offering the token or password are not necessary authentic. Biometrics authenticates the user directly.

If it is considered that a user generates a number of biometric tokens every day, each one unique, this method of encryption is closely related to one time pad technology – the keys used, are very long and do not form any pattern. As each key are used, this biometric key is marked as used for encryption by the server in the SBA, and will not be used ever again.

# 6 References

[1] Tait, B.L., von Solms, S.H. "Solving the problem of replay in Biometrics- An electronic commerce Example". Proceedings of 5th IFIP Conference on Challenges of expanding internet: E-commerce, E-business, and E-government. (I3E'2005) p468-479. Springer – ISBN 0-387-28753-1. Poznan, Poland 28-30 October 2005.

[2] James Wayman, Anil Jain, Davide Maltoni, Dario Maio, BiometricSystems: Technology, Design and Performance Evaluation, Springer; 1 edition (December 16, 2004), ISBN 978-852335960.

[3] Charles P. Pfleeger, Shari Lawrence Pfleeger, "Security in Computing" Third edition, Prentice Hall, ISBN 0-13-035548-8.

[4] Ebay online Auction: http://www.ebay.com / http://www.ebay.co.uk.

[5] PayPal online payment environment: http://www.paypal.com.

[6] Tait, B.L., von Solms, S.H., BioVault: a Secure Networked Biometric protocol, D.Com Dissertation, University of Johannesburg, 2008.

[7] Tait, B.L., von Solms, S.H. , Secure Biometrically Based Authentication Protocol for a Public Network Environment, Proceedings for the 4th International Conference on Global E-Security 23 – 25 June 2008, University of East-London, Docklands, United Kingdom, p238 – p246.

[8] Howard Wolinsky "Tagging products and people. Despite much controversy, radiofrequency identification chips have great potential" EGE (2005b) Ethical Aspects of ICT Implants in the Human Body MEMO/05/97, 17 March. Brussels, Belgium.

[9] EGE (2005b) Ethical Aspects of ICT Implants in the Human Body: Opinion Presented to the Commission by the European Group on Ethics. MEMO/05/97, 17 March. Brussels, Belgium: European Group on Ethics in Science and New Technologies.