USING A CENTRAL DATA REPOSITORY FOR BIOMETRIC AUTHENTICATION IN PASSPORT SYSTEMS[1]

Morné Breedt[a,b], **Martin S Olivier**[b]

[a]eyedentity

[b]Information and Computer Security Architectures Research Group,

Department of Computer Science, University of Pretoria, South Africa

ABSTRACT

Passports and visas are currently undergoing a rapid change due to legislation being passed by authorities such as the US requirement that they should have machine readable biometrics embedded within them. What the governments are trying to accomplish is to positively link individuals to these identity documents. But if an impostor is capable of forging a passport, complete with embedded biometric, the document could still pass the biometric authentication step. To prevent the system from being compromised in this manner, a central repository of biometric templates could be incorporated. Then the system could compare its 'live' biometric reading with the remote repository, rather than relying solely on the authenticity of the biometric embedded within the identity document.

Using a central repository raises certain questions. Firstly, how will the system know which repository to query? How can the system be sure the passport is not fraudulent and referring queries to a fake repository? How will the repository know it is communicating with a valid port system and not just wasting time with fraudulent queries performing a denial of service attack?

This paper proposes a model for the construction of a passport system that employs a set of national repositories for biometric identification that addresses these issues.

KEY WORDS

Biometric, Passport, PKI

[1] This material is based upon work supported by the National Research Foundation under Grant number 2054024 as well as by Telkom and IST through THRIP. Any opinion, findings and conclusions or recommendations expressed in this material are those of the authors and therefore the NRF, Telkom and IST do not accept any liability thereto.

1 INTRODUCTION

Most countries protect the integrity of their borders through strict access control. After the September 11, 2001 attack on the World Trade centre, American and other authorities have started to further increase their security and safety measures to prevent future terrorist attacks. Traditionally, access control into and out of a country has been the reserve of ID documents such as passports and visas. However, more and more countries are recognising that more is needed to positively identify an individual in a world where identity theft is rife and forged passports are commonplace. One of the steps countries are taking to improve their identity documents is the inclusion of machine readable biometrics — to help ensure the document does actually belong to the individual presenting it.

An example of this is the Enhanced Border Security and Visa Entry Reform Act of 2002 [1] - introduced in the US House of Representatives on December 19, 2001 which states:

"Not later than October 26, 2004, the government of each country that is designated to participate in the visa waiver program established under section 217 of the Immigration and Nationality Act shall certify, as a condition for designation or continuation of that designation, that it has a program to issue to its nationals machine readable passports that are tamper-resistant and incorporate biometric and document authentication identifiers that comply with applicable biometric and document identifying standards established by the International Civil Aviation Organization."

Currently the International Civil Aviation Organization (ICAO) is the international authority charged with the task of developing a standard for machine readable travel documents (MRTD). 1986 saw the formation of a Technical Advisory Group on Machine Readable Passports by the ICAO. The ICAO describes three types of MRTD [2]:

- A passport which indicates a person is a citizen of the issuing country;
- A visa used to indicate that the issuing country grants a non-citizen the rights to enter the country for a set period;
- Other travel documents which could be issued to non-citizens for travel across borders. An example would be a special purpose identification/border-crossing card.

To be able to incorporate a biometric in the above mentioned documents, the ICAO suggests three different implementation options [2]:

- The biometric will be contained within the document by making use of an appropriate storage medium. For example, a chip, magnetic stripe or 2D barcode.
- The biometric templates will be held by the issuing bodies (for example, a central database at each embassy in the case of foreign countries).
- The biometric will be extracted from a visual element within the document (for example, the photo of the face in the document).

While storing a biometric within the travel document - option 1 and 3 above – does allow authorities to 'biometrically' link the document to the bearer, this situation is still

open to potential compromise, namely, possible reproduction by an unauthorised individual or group. This is because the actual template is stored locally on the document. So, if a skilful forger can replicate the biometric capturing and storing process — be it through inside help, espionage, or other means — the production of fake identity documents would become possible. One can only imagine how easy this process would be if option 3 is used and the biometric template is being extracted from the photo in the passport.

Current chip technologies make it extremely difficult for an individual to tamper or forge the content on the chip. However, in the unlikely event the above scenario does occur, authorities will have to amend their current processes. This would require an extremely costly process of recalling and reissuing passports. For these reasons, the second option – a central biometric repository – would appear the most satisfactory since this will allow us with a secure environment (much like a secure chip/smartcard etc) as well as the flexibility to alter the biometric template in the event of a compromise.

Putting aside the issues of privacy and current legal requirements surrounding identity documents, the remainder of this article will consider the use of a central biometric repository for international border control. The use of a single international biometric authority appears impractical, not only because it will require countries to relinquish some of their autonomy, but also force every nation to conform to a biometric implementation their citizens might have rejected for sociological, technical, economical or political reasons.

Allowing countries to maintain their own national repositories seems to be the solution. However, this does present some problems. Firstly, the port system needs to know which national repository to query and prevent communication with a masquerading repository. Secondly, the biometric repository needs to be sure it is communicating with a valid port system. Thirdly, all communication needs to be safeguarded to prevent sample capture and replay, and also to prevent unlimited access to a large number of raw biometric samples (since these samples could be used to commit identity theft and espionage). And finally, for redundancy purposes, an alternative method needs to exist in the event that a national repository cannot be reached. While some of these issues are easy to address in a more general context, the autonomy and divergence of the countries participating in the process makes solving them more complex. This paper proposes a model that solves these issues at, as minimal an expense to national autonomy as possible.

The remainder of this article will be structured as follows: Section 2 will briefly introduce the field of biometric identification and outline a generic model for the function of any biometric system; Section 3 will contain a model for travel documents (MRTDs); Section 4 will examine the major implementation concerns when incorporating the model, and Section 5 contains the concluding remarks.

2 BIOMETRIC IDENTIFICATION

2.1 Background

Biometric authentication is considered the automatic identification, or identity verification, of an individual using either a biological feature they possess (physiological characteristic such as a fingerprint) or something they do (behavioural characteristic, such as a signature) [9]. Within each group there are also a wide variety of characteristics that

can be measured and, although the actual systems used are often very different, the basic processes and functions they perform can be described by a generic model, according to Wayman [10].

The model is divided into five sub-sections: data collection; transmission; signal processing; storage, and decision.

2.1.1 Data Collection

The data collection sub-system will be responsible for acquiring the biometric sample and transforming it into an electronic output. The system output depends on:

1. the biometric being measured;
2. the way in which the biometric is presented; and
3. the technical characteristics of the sensor used.

If any of the above dependencies change, the repeatability and the uniqueness of the biometric could be compromised.

2.1.2 Transmission

A number of the biometric systems available acquire their biometric samples in one location and process them in another. This remote processing allows them to centralise the administration of the system and reduce costs.

In such centralised biometric systems (where the processing/storage server is hosted at a different physical location) a transmission system is required. Moreover, if a great amount of data needs to be transmitted, a compression system will also be required. One last possible addition to the model described by Wayman is encryption before transmission — if we are going to make use of a public network for communication, for instance.

2.1.3 Signal processing

Once the biometric has been extracted by the sensor and, if needed, transmitted to the processing unit, it needs to be matched against the database. In order to do this the biometric sample — the gathered “image” (the data collection system output) — must be ‘prepared’. In Wayman’s model, this preparatory process is divided into three sections: quality control, feature extraction, and pattern matching.

Quality control is applied to the “image” gathered by the sensor so that the feature extraction process can acquire the true biometric pattern. In other words, a 'cleaning' process must be used to filter out the noise generated by the sensor and transmission of the “image”. Once 'cleaned', the system starts to search the pattern for unique, repeatable features. Once the suitable features have been identified, the system digitises them into a binary representation. This binary representation is usually considerably smaller than the “image” gathered by the sensor and cannot be reversed back into the original sample (it is very similar to one way hashing used for MD5 password protection — as used, for example, in Unix operating systems).

After the sample has been processed by the feature extractor, it is ready to be used in one of the two main functions of biometrics: enrolment or identification.

For both functions pattern matching is essential. The purpose of pattern matching is to compare the given sample to a number of templates in a database (the number of templates used will depend on whether it is performing identification or verification).

During this comparison, the pattern matching algorithm will determine how many features match and how many do not. This measurement will then be given to the decision sub-system. During enrolment, pattern matching will ensure that we are not enrolling a duplicate template into the database; during recognition, pattern matching will identify (one-to-many) or verify (one-to-one) the individual.

2.1.4 Decision

The decision making sub-system will receive data from the signal processing unit regarding the amount of non-matches or dissimilarities. The number of dissimilarities between a biometric sample and template is known as the Hamming Distance [14].

The decision sub-system will now determine a match or non-match according to the Hamming Distance computed, in conjunction with the policy of the system in question. The system policy will specify a 'cut-off' Hamming Distance — the threshold above which the system will reject the sample and a non-match will be deemed to have occurred. Conversely, if the Hamming Distance computed is lower than this prescribed threshold it will be deemed a match.

Many biometric systems allow the operators of the system to specify this cut-off threshold and this, in turn, could affect the accuracy of the biometric.

2.1.5 Storage Sub-system

Within biometric systems there are two different types of data that can be stored in the biometric database. The first is the hashed biometric code — produced after feature extraction and used in future recognitions. The hashed biometric code or template can be stored on a number of different media including a central database, smartcards, and magnetic strips. The second type is the raw biometric sample — gathered from the sensor in the data collection sub-system. There are two good motivations for storing the raw biometric data and not just the template: firstly, the ability to "re-issue" the biometric code quickly and easily (by running the raw biometric through the feature extraction process again), and, secondly, the ease with which the feature extraction and decision making phases can be modified (by simply running the raw biometric through the new algorithms). These advantages also allow the biometric system to be upgraded to a new version (or new vendor) easily and without the hassle of re-enrolling all the users.

2.2 National identification examples

2.2.1 Philippine Social Security System

On Nov. 17, 1998, the Philippine Social Security System (SSS) launched its ID card system [11]. The system is an AFIS (Automated Fingerprint Identification System) card system which will be used to ensure that members, pensioners and dependants do not enrol using multiple identities.

2.2.2 Biometrics in the electoral process

In a summary report for the National Biometric Test Centre [12], Wayman discusses the possibility of using a biometric in the electoral process to eliminate individuals from placing multiple votes.

Wayman suggests altering the voter registration and voting processes in the following way: During registration, individuals will be asked to supply a biometric sample along with their identification information. This sample could then be used both during registration to prevent an individual from registering using multiple identities and during the voting process itself to ensure that the person arriving at the voting station is a valid, registered voter. This latter identification could be accomplished by comparing a sample taken at the voting station to the enrolled templates stored in a centralized database managed by the voting jurisdiction.

3 THE MODEL

As mentioned previously, the two main phases or functions of any biometric system are: enrolment and identification. The model proposed in this paper focuses on the identification process and, in particular, border control — what is required to satisfactorily identify a person when they enter a country and present their machine readable travel document (MRTD).

The model for the recognition procedure can be viewed in Figure 1 (client side) and Figure 2 (server side) below. Five main 'areas' have been identified based on the functions they perform: the document read, biometric scan, client, server, and template database.

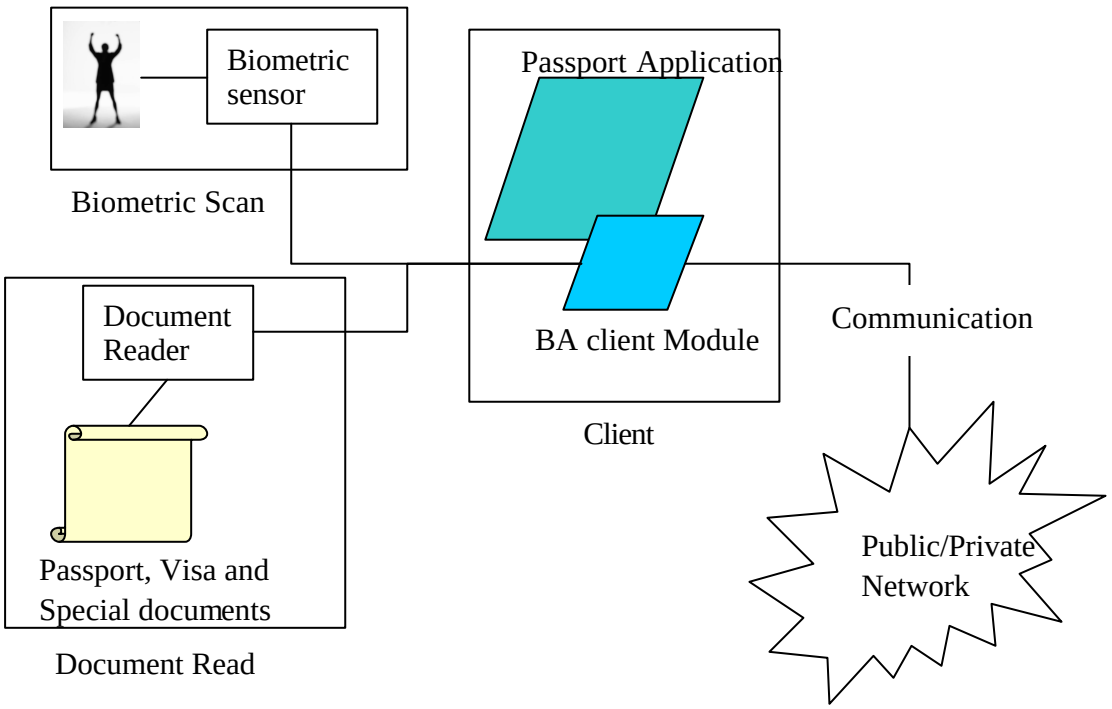

Figure 1 Client side of a biometric passport system

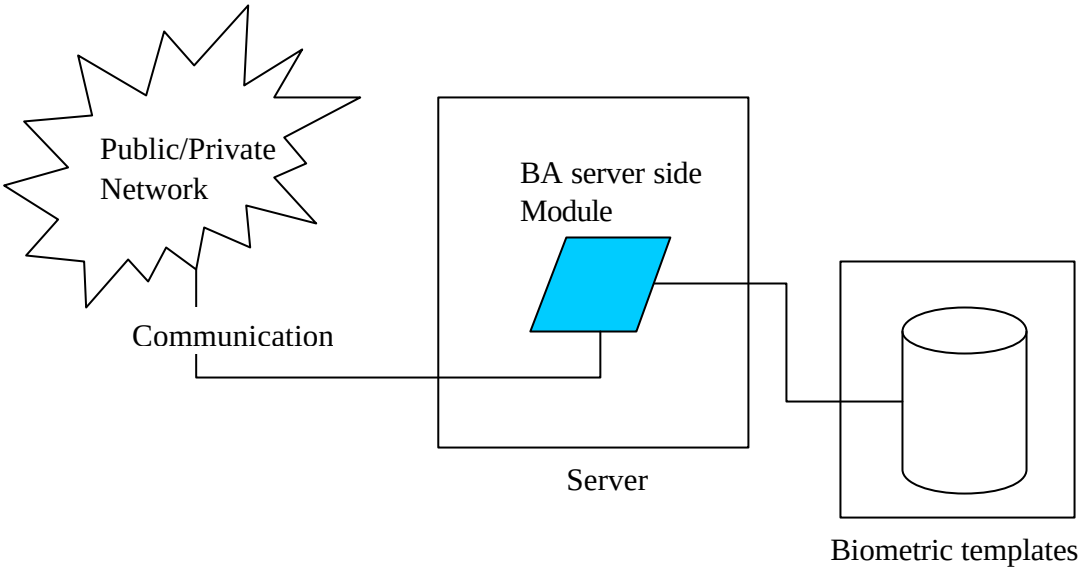

Figure 2 Server side of a passport system

3.1 Document Read

The information within the travel documents needs to be able to facilitate a number of different functions. Firstly, it needs to adequately identify the individual to whom the document belongs. Secondly, it needs to be able to identify which server to connect to (since there may be a number of different servers available, be it for load balancing or redundancy purposes). Thirdly, the authorities need to be able to perform an 'off-line' recognition if the connectivity of the border center goes down. And finally, it needs to ensure the integrity of the data contained within the document.

Two main components are required to accomplish the above: a document certificate (see 3.1.1), and a biometric template.

Storing the biometric template within the document will solve the problem of off-line recognitions in the event that the central biometric repository is unavailable. For the purposes of this paper, a detailed description of the template is not required since that will be specific to the type of biometric and vendor chosen by the issuing authority. More important, from the point of view of the model, is the document certificate and its usage.

3.1.1 Document Certificate

A document certificate (DocC) is an electronic certificate similar to those used in electronic commerce. A large amount and wide variety of different types of information can be stored within the document certificate including information about the individual (passport number, name etc), and the public key from a private/public key pair. This key pair will be generated at the production of each travel document by the issuing authority. The issuing authority will store the public key within the document and the private key with the biometric template. A third piece of data the DocC will contain is identification information relating to the data store(s) being used (multiple data stores are likely to be used to ensure adequate redundancy). In other words this data will point the port-system to the appropriate data store. And the last piece of data in the DocC will be supporting information such as version numbers of the biometric software.

All the information in the DocC will, as with traditional certificates, be signed by creating a hash of the information and then encrypting the hash with the issuing authority's private key [8]. The signing and encryption of the DocC by the issuing authority now requires the client module to have a copy of the issuing authority's public key in order to verify the signature. A copy of the key can be obtained in a number of different ways, for example, published on a public server which is made safe or trustworthy by using certificate authorities and different trust models [7] — see below for further considerations regarding incorporating a PKI trust model.

The signed DocC addresses a large number of the issues mentioned previously and the system is now capable of verifying the legitimacy of the issuing source, accurately identifying the individual and ascertaining which server to use. This is possible because contained within each certificate is the public key of the issuing authority — the only key capable of decrypting the hashed DocC — proving that the DocC is from a valid source (verifying the signature) [8]. Furthermore, the signing of the document guarantees that the information has not been compromised (altered or replaced) because, once the document has been decrypted, the hash can be re-used on the information in the DocC and compared to the originally decrypted hash. If the two hashes match, the information has not been altered.

3.2 Biometric Scan

As mentioned previously, the data collection process makes use of scanners to capture a biometric sample and transform it into an electronic signal. For this application, the scanner can be anything from an optical fingerprint scanner to a normal web camera (used during face recognition). Although the model is flexible regarding the scanner used, the system still needs to make sure that the scanner is an authorised scanner and not a replacement generating fake readings. Consequently, some form of identification mechanism for the scanners is required. A possible solution would be to issue each scanner with its own digital certificate and embed this certificate within the scanner. This would allow the client module to verify the authenticity of the scanner before accepting a biometric sample from it. Moreover, the certificate cannot only be used to positively identify the sensor, but to ensure the integrity of the biometric sample as well. By signing the raw biometric sample using the certificate, the client or server module can verify that the sample is from a legitimate source and establish trust that the sample has not been altered since it left the inner sanctity of the scanner (data integrity).

Further benefits of this solution include the safe incorporation of additional cryptography techniques to prevent unauthorised access to the raw biometric samples and ensure that a sample is not replayable. For example, signing the template with a random key that will only be valid for a limited time and only be used once (the key could be safely transmitted to the sensor by encrypting it with the sensor's public key).

3.3 Server

The server side of the passport biometric authority consists out of two parts. Firstly, a biometric template database which will contain the templates of all the individuals enrolled on the BA. The second is the server side module which is responsible for a number of the operations described in Wayman's generic biometric model [10].

The first operation to be performed by the server will be 'signal processing' or the feature extraction process. The reason for suggesting that the feature extraction process resides on the server is updatability. This allows the system's administrator to update the feature extraction process (when new releases become available, for instance) with minimal impact on the system — since it will be restricted to the issuing authorities' servers and not all the border posts' systems.

Once the server has generated a feature template it can begin its second function: decision making. For the purposes of this particular application, the decision making process could be accomplished by performing verification (one-to-one match). The advantage of performing a verification instead of a recognition (one-to-many) at this stage is the speed of the decision making process. Table 1 gives an indication of the speeds of different matching algorithms [13].

Biometric	**Matches per Minute**
Face	800
Fingerprint chip	60
Fingerprint chip (2)	2 500
Fingerprint optical	50
Hand	80 000
Iris	1 500 000
Voice	680

Table 1 Matching speeds of different biometric algorithms [13]

For verification, the stored template will be extracted from the template database by making use of the information stored within the individual's DocC. The results of the verification will be submitted back to the client for processing.

3.4 Client

In order to facilitate authentication the client module needs to perform the following functions of the generic model: data collection and transmission. In this model, the data collected will include the biometric scan and the data contained within the document. And once the client has the data it will be transmitted to the server for processing.

One important point to note is that the feature extraction phase (discussed in the server side) could be moved to the client. Thus the features will be extracted at the client and a "feature packet" will be sent to the server for matching.

3.4.1 Allowing multiple biometrics in one client terminal

For one application to be capable of using multiple biometrics the biometric functions need to be separated from the functions of the application. To a degree, this has already been accomplished in the above model since most of the biometric sample processing has been moved to the server. The only biometric related functions the client actually needs to perform are controlling the sensor, receiving and then transmitting the raw sample to the server for processing.

To accomplish this, the client has been divided into two separate components: the client application — which will be the port system, and client module — which will

handle all the client aspects of the biometric process and return the results of the process to the client application.

This approach creates two main benefits: the ability to alter or update the biometric process on the client by simply replacing the client module and not the entire application and, secondly (and more importantly), the ability to use more than one client module (enabling the use of multiple biometrics).

But why permit the use of multiple biometrics? Using the USA as an example, its Enhanced Border Security and Visa Entry Reform Act requires their passports, visa waiver country's passports and the US visas for non-visa waiver passports to contain a machine readable biometric [1]. For their own visas and passports the US can prescribe which biometric they want to use (a single biometric), but the visa waiver countries might want to use another biometric for social, cultural (public opinion will have a large impact on the success of a biometric system) or economical reasons. If they opt for a different biometric it would leave America in a predicament. America can simply revoke their visa waiver status and require them to use a visa, but America might want to have a good international relationship with the country in question and consequently have to try and incorporate the other country's system into theirs. However, by adopting the multiple client module approach, America's developers can develop one port system that calls the appropriate module depending on the visitor and simply processes the response.

One last point to bear in mind is that the generation of passport and other travel documents is controlled by international standards regarding facets such as the storage medium. Therefore, to incorporate a multi-biometric passport application we could only need one document reader and a set of different biometric readers with an appropriate client module for each (hopefully biometric standards will allow us to use one scanner per biometric type and not end up using, for example, 10 different fingerprint scanners because 10 countries decided to use different vendors) .

3.5 Template Data Store

In *Hiding Biometric Data* [3] Jain and Uludag list a number of possible attacks a biometric system can face. Within this list there are two possible attacks relating to the template database. Firstly, the alteration of the templates on the data store, and, secondly, altering the template while being communicated from the data store to the matching module. The implication of any one of the above attacks is that the matching process can now be fooled into giving a positive result. For this reason it is crucial that the biometric template be stored and transmitted in encrypted format, be it asymmetric or symmetric encryption.

4 IMPLEMENTATION CONCERNS

An important requirement for the port system is to know it is communicating with a valid server, and also for the server to know it is communicating with a valid port-system. One reason for this is to ensure that the server is not wasting time with transmissions from a pseudo port system which is 1) trying a denial of service attack and/or 2) analysing multiple responses to determine how it can intercept and alter a response. The client on the other hand needs to be sure it is communicating with a valid server to prevent it from accepting a response from a pseudo server which is simply giving positive responses for its fraudulent travel documents in circulation.

To facilitate the identification required we can make use of the digital certificates issued to the client and server mentioned in the discussion above. The use of such certificates is now pretty straightforward and has been used with great success in a number of applications in fields such as e-commerce. But critical to the success of a process using such certificates is the trust we can place in the certificates.

4.1 Trust relations

Traditionally, the authenticity of certificates has been validated by having the certificate digitally signed by a known (and trusted) entity, and by doing so stating that the certificate and owner are valid. These entities are known as Certificate Authorities (CAs) [7, 8]. A number of different implementation models exist [7], for example, a single CA model. In such a case there would be a single "vouching institute" for the certificates; an easy pick for this global CA would be the United Nations. But, as mentioned by Perlman [7], there is not one global organisation trusted by all countries, corporations, religious groups, political groups, etc. The single CA approach could thus turn out to be impractical, especially when so many different countries start to play in the same area. Consequently, a multiple CA approach consisting out of CAs that are trusted by the countries involved appears the most practical solution. An example of this approach is the Configure plus delegate CAs approach [7] currently being used in browsers.

5 CONCLUSION

The threat of terrorism has never been so prevalent, and this precipitated a global mind shift about the way security is implemented. Countries are looking for alternative ways of increasing protection and tamper-proof biometric travel documents have come to the fore as a viable budgetary consideration.

Of all the different ways of incorporating a biometric within a machine readable document, the most promising seems to be the use of a biometric authority (central database). Although this in itself does not make the system foolproof it does raise the cost of forgery beyond what would be viable for most forgers.

The above article describes how the use of a central data repository which incorporates PKI extensively could raise the cost of forgery to a prohibitive level and thus prevent fraudulent travel documents from being circulated. Although this article focuses predominantly upon the recognition process, it is equally important that, when the document is being produced, the architects incorporate a biometric enrolment process to prevent an individual from registering for a travel document by making use of multiple identities much like the process involved in [11] and [12]. One will also have to consider the significant privacy issues faced when using a central national repository. One option is to consider using a distributed repository, where biometric details of different individuals of a single country are stored in different repositories spread over that country, or even where a single individual's biometric information is fragmented over more than one repository. This, however, is left for future research.

REFERENCES:

[1] USA, **Enhanced Border Security and Visa Entry Reform Act of 2002** PUBLIC LAW 107–173 MAY 14, 2002 (116 STAT. 554)

[2] ICAO TAG MRTD/NTWG, **Biometric deployment of machine readable travel documents**, Technical Report Version 1.9 19, May 2003

[3] Anil K. Jain and Umut Uludag, **Hiding Biometric Data,** IEEE Transactions on pattern analysis and machine intelligence, **25**, 11, 2003, pages 1494 – 1498

[4] Charles P. Pfleeger, **Security in computing, S**econd edition, Prentice Hall, 1996

[5] Mike Ellis, **Use of Contactless Integrated Circuits In Machine Readable Travel Documents**, Technical Report Version 3.1, WG3 for NTWG, 16 April 2003

[6] William A. Shay, **Understanding data communication & networks,** Second edition, Thomson, 1999

[7] Radia Perlman, **An overview of PKI trust models,** IEEE Network, **13**, 6, 1999, pages 38 – 43

[8] Sebastiaan H. von Solms, Jan H.P. Eloff, Mariki Eloff, and Elme Smith, **Information security,** Amabhuku, 2000.

[9] James L. Wayman and Lisa Alyea, **Picking the Best Biometric for Your Applications**, National Biometric Test Center Collected Works, vol. 1, J. L. Wayman (Ed.), San Jose, CA: National Biometric Test Center, 2000, pages 269 – 275

[10] James L. Wayman, **Fundamentals of Biometric Authentication Technologies** National Biometric Test Center Collected Works, vol. 1, J. L. Wayman (Ed.), San Jose, CA: National Biometric Test Center, 2000, pages 1 – 19

[11] James L. Wayman, **Philippine Social Security System Inaugurates Huge Civilian ID Card/AFIS System**, National Biometric Test Center Collected Works, vol. 1, J. L. Wayman (Ed.), San Jose, CA: National Biometric Test Center, 2000, pages 169 – 171 (Originally published in "Biometrics In Human Services User Group Newsletter" #12, 1999)

[12] James L. Wayman, **Biometric Identification Technologies in Election Processes : Summary Report**, National Biometric Test Center Collected Works, vol. 1, J. L. Wayman (Ed.), San Jose, CA: National Biometric Test Center, 2000, pages 253 – 261

[13] Tony Mansfield, Gavin Kelly, David Chandler and Jan Kane, **Biometric Product Testing Final Report**, Centre for Mathematics and Scientific Computing, National Physical Laboratory, Issue 1.0, 19 March 2001

[14] John Daugman, **High Confidence Visual Recognition of persons by a Test of Statistical Independence**, IEEE Transactions on Pattern analysis and machine intelligence, **15**, 11, 1993, pages 1148-1160.